

1 CCTGGAAGTG CCAGGGAGCA CTGGAGGCCA CCCAGTCATG GGGGACACCT
 51 TCATCCGTCA CATCGCCCTG CTGGGCTTTG AGAACGGCTT CGTACCCAGC
 101 CAGCACTATG TGACATGTT CCTGGTAAA TGGCAGGACC TGTGGAGAA
 151 GGTGGTCTAC CGGCCTTCA CCGAGATCTA CGAGTCCAT AAAACCTTAA
 201 AAGAAATGTT CCCTATTGAG GCAGGGGCGA TCAATCCAGA GAACAGGATC
 251 ATCCCCCACC TCCCAGCTCC CAAGTGGTT GACGGGCAGC GGCCGCCGA
 301 GAACCGCCAG GGCACACTTA CCGAGTACTG CAGCACGCTC ATGAGCCTGC
 351 CCACCAAGAT CTCCCGCTGT CCCCCACCTCC TCGACTTCTT CAAGGTGCGC
 401 CCTGATGACC TCAAGCTCCC CACGGACAAC CAGACAAAAA AGCCAGAGAC
 451 ATACTTGATG CCCAAAGATG GCAAGAGTAC CGGCACAGAC ATCACGGCC
 501 CCATCATCCT GCAGACGTAC CGGCCATCTG CCAACTACGA GAAGACCTCG
 551 GGCTCCGAGA TGCGTCTGTC CACGGGGGAC GTGGTGGAGG TCGTAGAGAA
 601 GAGCGAGAGC GGTGGTGGT TCTGTCAGAT GAAAGCAAAG CGAGGCTGGA
 651 TCCCAGCGTC CTTCCTCGAG CCCCCGGACA GTCTGACGA GACGGAAGAC
 701 CCTGAGCCA ACTATGCAGG TGAGCCATAC GTGCCATCA AGGCCTACAC
 751 TGCTGTGGAG GGGGACGAGG TGTCCTCTGCT CGAGGGTGA GCTGTTGAGG
 801 TCATTCAAA GCTCTGGAC GGCTGAAAG ACGACGTCAC AGGCTACTTC
 851 CCGTCCATGT ACCTGCAAAA GTCAGGGAA GACGTGTCAGC AGGCCAACG
 901 CCAGATCAAG CGGGGGGCGC CGCCCCCGCAG GTCTCCATC CGAACCGCGC
 951 ACAGCATCCA CCAGCGGTGCG CGGAAGCCGC TCAGCCAGGA CGCCTATCGC
 1001 CGAACACAGG TCCGTTTCT GCAGCAGCGA CGCCGCCAGG CGGGGCCGGG
 1051 ACCGAGAGC CCGGGGAGCC CGCTCGAGGA GGAGCGGCAG ACCGAGCGCT
 1101 CTAAACCGCA CGGGGGGGTG CCCCCCGCCG CGAGCGCCGA CCTCATCTG
 1151 AACCGCTGCA GCGAGAGCAC CAAGCGGAAG CTGGCGTCTG CCGTCTGAGG
 1201 CTGGAGCGCA GTCCCCAGCT AGCGTCTCGG CCCTTGGCGC CCCGTGCGTG
 1251 TATATACGTG TTCTATAGAG CCTGGCGTCT GGACGGCGAG GGAGCCCCG
 1301 ACCCTGTCC AGGGGGCTC CCGCCACCT CAATAATGT TGCTTGGAGT
 1351 GGAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA

(SEQ ID NO: 1)

FEATURES:

5'UTR: 1 - 37
 Start Codon: 38
 Stop Codon: 1196
 3'UTR: 1199

Homologous proteins:

Top 10 BLAST Hits:

Sequences producing significant alignments:	Score (bits)	E value
CRA 18000004925255 /altid=gi 4557785 /def=ref NP_000256.1 neut...	789	0.0
CRA 18000005124568 /altid=gi 2754713 /def=gb AAB95193.1 (U5783...	788	0.0
CRA 18000005207006 /altid=gi 4263750 /def=gb AAD15422.1 (AC004...	783	0.0
CRA 18000005171728 /altid=gi 6685673 /def=sp 077774 NCFL_BOVIN ...	684	0.0
CRA 148000004473069 /altid=gi 8439513 /def=dbj BAA96544.1 (AB0...	670	0.0
CRA 118000005118410 /altid=gi 9623382 /def=gb AAF90134.1 AF2677...	663	0.0
CRA 18000005141875 /altid=gi 3061282 /def=dbj BAA25649.1 (AB00...	659	0.0
CRA 18000005020732 /altid=gi 1171669 /def=sp Q09014 NCFL_MOUSE ...	655	0.0
CRA 18000004937799 /altid=gi 2118398 /def=pir I54525_leukemia-...	651	0.0
CRA 148000001425618 /altid=gi 7839599 /def=gb AAF70344.1 (AF26...	527	e-148

EST:

Sequences producing significant alignments:	Score (bits)	E value
gi 12896059 /dataset=dbest /taxon=960...	1532	0.0
gi 12951967 /dataset=dbest /taxon=960...	1501	0.0
gi 12342004 /dataset=dbest /taxon=96...	1423	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

gi|12896059 placenta
 gi|12951967 B cells from Burkitt lymphoma
 gi|12342004 primary B-cells from tonsils

Tissue Expression:

Leukocyte

FIGURE 1

1 MGDTFIRHIA LLGFEKRFVP SQHYVYMLV KWQDLSEKW YRRFTEIYEF
51 HKTLMEMFPI EAGAINPENR IIPHLPAKRW FDGQRAAENR QGTLTEYCST
101 LMSLPTKISR CPHILDFFKV RPDDKLPLTD NQTKKPETYL MPKDGGSTAT
151 DITGPIILQQT YRAIANYEKT SGSEMASTG DWEVWEKSE SGWFCQMK
201 KRGWIPASFL EPLDSPDETE DPEPNYAGEP YVAIKAYTAV EGDEVSLLLEG
251 EADEVIHKLL DGWKKDDVTGY FPSMYLQKSG QDVSQAQRQI KRGAPPRRSS
301 IRNAHSIHQR SRKRLSQDAY RRNSVRFLQQ RRRQARPGPQ SPGSPLLEER
351 QTQRSKPQPA VPPRPSADLI LNRCSESTKR KLASAV
(SEQ ID NO: 2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION

N-glycosylation site

131-134 NQTK

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE

CAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 4

1	42-45	RRFT
2	297-300	RRSS
3	313-316	KRLS
4	321-324	RRNS

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 10

1	36-38	SEK
2	53-55	TLK
3	133-135	TKK
4	160-162	TYR
5	300-302	SIR
6	311-313	SRK
7	324-326	SVR
8	352-354	TQR
9	377-379	STK
10	378-380	TKR

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 11

1	53-56	TLKE
2	93-96	TLTE
3	148-151	TATD
4	171-174	SGSE
5	178-181	STGD
6	208-211	SFLE
7	215-218	SPDE
8	238-241	TAVE
9	246-249	SLLE
10	279-282	SGQD
11	344-347	SPLE

FIGURE 2A

[5] PDOC00008 PS00008 MYRISTYL
N-myristylation site

Number of matches: 3

1 83-88 GQRAAE
2 172-177 GSEMAL
3 280-285 QDVSQ

Membrane spanning structure and domains:

NO DATA

BLAST Alignment to Top Hit:

>CRA|18000004925255 /altid=gi|4557785 /def=ref|NP_000256.1| neutrophil cytosolic factor 1; Neutrophil cytosolic factor-1 (47kD); p47phox [Homo sapiens] /org=Homo sapiens /taxon=9606 /dataset=nraa /length=390
Length = 390

Score = 789 bits (2015), Expect = 0.0
Identities = 385/390 (98%), Positives = 386/390 (98%), Gaps = 4/390 (1%)
Frame = +2

Query: 38	MGDTFIRHIALLGFEKRFVPSQHYYMFLVKWQDLSEKWYRRFTEIYEFHKTLCMFPI	217
Sbjct: 1	MGDTFIRHIALLGFEKRFVPSQHYYMFLVKWQDLSEKWYRRFTEIYEFHKTLCMFPI	60
Query: 218	EAGAINPENRIIPHLPAPKWFDGQRAAENRQGTLTEYCSTLMSLPTKISRCPHLLDFKV	397
Sbjct: 61	EAGAINPENRIIPHLPAPKWFDGQRAAENRQGTLTEYCSTLMSLPTKISRCPHLLDFKV	120
Query: 398	RPDQLKLPTDNQTKKPETYLMPKDGSSTATDTGPIILQTYRAIA+YEKTSSEMALSTG	577
Sbjct: 121	RPDQLKLPTDNQTKKPETYLMPKDGSSTATDTGPIILQTYRAIA+YEKTSSEMALSTG	180
Query: 578	DWEWEKSESQWWFCQMKAKRGWIPASFLEPLDSPDETEDPEPNYAGEPYVAIKAYTAV	757
Sbjct: 181	DWEWEKSESQWWFCQMKAKRGWIPASFLEPLDSPDETEDPEPNYAGEPYVAIKAYTAV	240
Query: 758	EGDEVSLLEGEAVEVIHKLLDGW----KDDVTGYFPMSYLQKSGQDVSAQRQIKRGAPP	925
Sbjct: 241	EGDEVSLLEGEAVEVIHKLLDGWVIRKDDVTGYFPMSYLQKSGQDVSAQRQIKRGAPP	300
Query: 926	RRSSIRNAHSIHQRSRKRLSQDAYRRNSVRFLQQRQQARPGPQSPGSPLEERQTQSK	1105
Sbjct: 301	RRSSIRNAHSIHQRSRKRLSQDAYRRNSVRFLQQRQQARPGPQSPGSPLEERQTQSK	360
Query: 1106	PQPAVPPRPSADLILNRCSESTKRKLASAV	1195
Sbjct: 361	PQPAVPPRPSADLILNRCSESTKRKLASAV	390 (SEQ ID NO: 4)

FIGURE 2B

Hmmer search results (Pfam):

HMM results:

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
CE00053	CE00053 mox_mitogenic_oxidase	573.6	1.3e-168	1
PF00787	PX domain	119.4	6.6e-32	1
PF00018	SH3 domain	107.5	2.7e-28	2
CE00036	CE00036 androstane_receptor	0.3	4.6	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmmer-f	hmmer-t	score	E-value
PF00787	1/1	4	121	..	1	147	119.4
PF00018	1/2	159	213	..	1	57	66.7
CE00036	1/1	243	257	..	199	213	0.3
PF00018	2/2	229	279	..	1	57	41.0
CE00053	1/1	1	386	[]	1	566	573.6

FIGURE 2C

1 TACTAAAAAT ACAAAATTAG CCAGGCGTGG TGGCGCACAC CTGTAATCCC
 51 AGCTACTTGG GAAGCTGAGG CAGGAGAAC GCTTGAACCT GGAAGGCAGA
 101 GGTTGCAGTG AGCCGAGATT GTGCCACTGC ACTCCAGCCT GGGCAACAAG
 151 AGCGAAACTT CGCTTCAAAC AAATAAAATTA ACGCCCCAGCA TGTCCTGGCT
 201 TTCATCTGCC AGACCTCAAC CCTCACCCCC AGGAGATCAG GTCCGGACCA
 251 TGAGCTGACC CTGGACTCAG GCAAGGGTGA GTTGGTGCAG CCCTGGCTG
 301 CTGGGAGGCA CAGGCTGCAG CAGGCTGCT GGGGCTGAGG CCCGCCACTC
 351 ATGAACCAT GACCTTGAAT GAGCTTAAAAG AGCTCTGGC CTCCCAGGCT
 401 CTAGGGGGAG TGGGAGAGAG AGGCCTCAGC CTGTCCTGG GCATGCTGCC
 451 CCCTCCTCAC CTCTTTGTCC CAAATCCCC TCCCTGGCAAA GCTGACAGTC
 501 TTAATATCAC TCTGGAGAAA ACTGAGTCAG CCCTAAGGAA CAATTCAATG
 551 AACCATTTGC TTACTTGGG ATTGGAACTC AAGTCTCACT CAAAGTCTGT
 601 GCCATTTCG TCCCAGCTGT CACTGGCCCT CATCCACACCA CACCCAGGA
 651 TGAGCATCTA ACGCTTGCAT GCACACTCCC ATGCCCGGT TCATTCACTC
 701 ATTCAATTATC TCATTCACTC ATTCAATTGAC TCATTCAATT ATTCACTCAC
 751 TCATTCAATT ACTCAGTGAA TGTTGCAGTC ACGATCCAAA TATTTATGGC
 801 CTCTGTGTC CAGGCACTAG ATGGAGGGC TGgggCTAGA GCCCCTGATA
 851 ACCCGGTAT GCCCTAGCTT TCCCTGGACA CACATTGGG TAAGGGGAGA
 901 CTAAAAAAAT TAAGTCAGGC CAGGCACCGT GGCTCATGCC TGAATCCAG
 951 CACTTTGGGAGGCCGG AGTGAATTAC CTGAGGTAG GAGTTCAAGA
 1001 CCAGGCTGGC CAACATGGAG AAACCCAGTC TCTAATTAAA AAAAAAA
 1051 AAATTAGCCA GGTGTGGTGG CACATGCCG TAATCCCAGC TACTCAGGAG
 1101 ACTAACGCAA GAGAATTGCT TGAACCCAGG AGGCAGAGGT TGCCGTGAGC
 1151 CGAGATCGCG CCATTGCACT CCAGCCTGG AAACAAGAGC GAGACTCCAT
 1201 CTAAAAAAA AAAAAAGTGG GAGGCAGAGG CAGGAGGATC ACTAGAGGCC
 1251 AGTAGTTGA GACCATCTG GGCAACATAG CAGGACCCCTG TCTGTACAAA
 1301 AAAATTAAA AAAATTAAAC CGGGCATGGT GGCACACACC CGTAGTCCA
 1351 GCTACTCCAG AGGCTGAGGC AGGAGGATCG CTGGAGCCCA GGAGTTGGAG
 1401 GCTGCAGTGA ACTGTGATCC CACCACTGCA CTTAACCTG GATAACAAAG
 1451 CAAGACCTG TCTCAAATAA CAATAGCAAT AATAATAAAAG AAAAATTAAA
 1501 TGCAATTGC GATGCATCAG TGATAAGTGC TCTGCAGAAA AAGGAGGCAG
 1551 GAAGAGGCTG AGAAAGGTAT GAGGTTTGTG ATGCAATGTG AAGTTATCAA
 1601 GGAAGGCTTC TCGGAAGAGG TGACATTGAG GCAAGAGAAAT GGAGGGAGAT
 1651 TATGGAGGGAGATGGTGGAA TGGGGGGAAAC ATGGTCAAGA CCAGGAATAT
 1701 GGTCAAGGGG GGAAGAGATGG TCAAGGGAC GCAGCAAATG CAAAGGCCCT
 1751 GAGGCAGGAG CAGCTTGATT CACCCCCAAA ACCCGTGGGG CCCGTGCAGG
 1801 CGAGGGGAAG GACAAGTGTA AACCCCTTTTC CTTGTCCTG CAGGTGTGTG
 1851 TGAACATGAG TCTGCCCATG TTTACACCT GCAAGCCTGA AGAGTCCCCA
 1901 GAAACTGAAA GAAGAAGCAA AGCCCTTCT GTACCCCTCCC TGCCCCCTGT
 1951 CCCGACCGG ACAAAAGCGA CTTCCCTTT CCAGTCATT TAAGGGCAG
 2001 CCTGAAAGTG CCAGGGAGCA CTGGAGGCCA CCCAGTCATG GGGGACACCT
 2051 TCATCCGTCA CATGCCCTG CTGGGCTTTG AGAAGCCTT CGTACCCAGC
 2101 CAGCATTGAG TGAGTAGCTG GTGGAGGGCA TCCCGTGGGG GGAATACGGG
 2151 AGGGACAGCA CGGGCCACCCCT TGCACTCCA GGGCCAAACCA GCTCCAGTGA
 2201 GGACTAACGG GGCAAGGGTCT TGGGCACCTG GTCCCTGGTC TTTGAGCTG
 2251 GATCTACCCC TCTGATCCCT GGGAAAGACAG TCCCTTGGG CCGCCCTGG
 2301 GCCCCAGCCC TTTACTGTCC CCGCCTGTGT CCCCAGCCAG GCCCTCAGCC
 2351 TTAGGCCAGGA GTCCCTTTTC TGCTCCCCCTG CCATGGCCAG GCAGGCCAGC
 2401 GCTCTCTCAG GTCCGAGGCC CACTCCCTCA GGAAGCCTTC CCTGACTAGC
 2451 CCAGCTATCA GAGAGTGGCC CTCCCAAGAG GGAGGCTGG AACTAAAGC
 2501 TCTCTCTCTC CCCAGCTGCC TGAGTAGTCA GTTAGAGTCT TATCCTCTCC
 2551 AGTAGGGTGA CACCATGACA GGGGCCAATA GAGTCCTCCC ATCTGTCCCC
 2601 AAGGAGGCTG GACAAATGCC TGCTCAGACA CACAAGTCCA CTGGGTCCCC
 2651 TAATCCATA GGAAGGCCAG GGAGGAACCA CATTAGGAA ATTGAAGCTT
 2701 GTATGGAACA TTTAGTCCCTA TGTCAGGAA CTTTCTCTT TTTTGTATT
 2751 TTTTTGTGT TTGAGACAGA GTCTTGATCT GTGCCCAGG CCAGAGTGC
 2801 GTGGCACGAT CTCACTCAC TGCAACCTCC GCCTTCCAGG TTCAACTGGT
 2851 TCTCTGCCCT CAGCCTCCAG AGTAGTTGG ATTACAGGTG CCCACCCACCA
 2901 CGCCTGGCTA ATTTCCTGAT TTTTAGTAGA GACAGGGTTT CACCATGTTG
 2951 GCCAGACTGG TCTCAAACCTC CTGACCTCAA GTGATCCACC CACCTGGGCC
 3001 TCCCCAAAGTG CTGGGATTAC AGGCATGAGC CACCGTGCCT GGCCTGTTT
 3051 TTGAAATGA GGTCTGGAGT GCAGTGGTGC GATCATAGTT CACTGCAGCC
 3101 TCAACCTCCC AGGCCAACAGT GATCCTCTG CTCAGCCCC TTGAGTAGCT
 3151 GGGGTACAG GGGCACACCA CCATGCCCTGG CTAGTTTTA AAATTTTGT
 3201 GGAGATGAGG TTTCACTATG TTGTCCAGGC TAATCTTGAA CTCCCTGGCT

FIGURE 3A

3251 TAAGCAACCC TCTGGTCTCA GCCTCCCACA GTGCTAGGAT TACAAGCGTG
 3301 AGCTACCGTG CCTAGTCACT TTTCTCCCTT TCTTTGTAAC TTTCAGTTTT
 3351 GAAATTTCAA ATTACAGAA AGGCTACTGG GTGTAAAAC GGTCAGCTC
 3401 ACTCCAATAG TCTTTCACTC ACCTTCATCC ACACCTCTCT TTCTGGGGAT
 3451 ATTTTCTGAA TTATTTGAGA GTGAGTTGAA GACGTGTTT CTTACCTCTA
 3501 AATACTAGTT GTTGGGCATT TCTTAAATC AAGGCATTCT CTTACATAAT
 3551 CACAACACAC GTGTCAAAAT CAGGAAATTA ACATGGACAA AACACCATA
 3601 TCCACCCACA GACTTTACTG AGGTTTCCCC GATTATCCTG CTTGTCCCT
 3651 GCAGTGAAA CTTTTTCAG GTCTAGGATC CAGTCAGGA TCAATGTCAT
 3701 AGCCTTAAAC CTTCCTTAAAT CTGGATCAGT CTTTTTCTT TTTCCTTTC
 3751 TTTTTTGGA CACGGAATCT CACTCTGTG CCAGACTGGA GTGCAGTGGT
 3801 GCAATCTCGG CTCTTGCAA CCTCTGCCCT CTGGGTTCAA GAGATTCTCC
 3851 TGCCCTAGCC TCCTAGTAG AGGTGCGCGC CACCACGCC
 3901 AGCTCGTTT TGGTAGAGAC AGGGTTTGC CATTGATTCT GGATCAGTCT
 3951 TTTTTTTTTT TTTTATGAGA TGGAGTCTTA CTCTGTCACC CAGGCTGGAG
 4001 TGCAATGGCA CAATCTCCAC TCACTGCATC CTCCGCCCTC CAGGTTCAAG
 4051 CAATTCTCGT GCCTCAGCCT CCCGAGTAGC TGGGATTACA GGATGCGCC
 4101 ACCATGCCCG GCTACTTTTT GTATTTTTAG TAGAGACAGG GTTTACCAT
 4151 GTTACGCCAGG CTGATCTCGA ACTCCCTGACG TCAGGTGATC TGCCCGCCTC
 4201 GACCTCCCAA AGTGCTGGG TTACAGGGGT GAGCCACCGT GCCAGCGGAT
 4251 TCTGGATCGG TCTTAATCAG TCTTTGTCTT TTGCAACTTT GATGTTTGC
 4301 AGAGAGCAGA CCAGTTACCT TGTAGAATGT CCCTTAGTTT GGGTTTATCT
 4351 TCATTAAGATT CAGTTTGTT ATCCAGGGCA GTGGATCTTA GATGCAATT
 4401 TGTCTTCTTT TTAATTTTTT TGAGAGGGAG TCTCGCTCTG TCAACCCAGGC
 4451 TGGAGTGCAG TGGCACAAAC TCAGCTCACT GCAGCCTCCG CCTCCCGGGT
 4501 TCAAGCAATT CTCTGTCCCT AGCCTCCCAA GTAGCTGGGA TCAACAGGTGC
 4551 CCATCACCAAC TACCGGGTAA TTTTTGTGTT TTTAGTAGAG ACAGGGTTTC
 4601 ACCATATTGG TCAGGCTGGT CTTGAACGCC TGACCTCAGG TGATCCACCT
 4651 GCCTTGGCCT CCCAAAGTGC TGGGATTACA GACGGGAGCC AACATGCCA
 4701 GCCTTCCCTGC CCCTCCCGTC CCCTCCCTC TCCTCTGTG CCCTCCCTC
 4751 CCCTCCCCTA TCCTCATGTC CCCTCCCTC CCCTCCCTC CCCACCCAAG
 4801 CTGGAGTGCA GTGGTGCAAT CATAGCTCAC TAAAGCCTTG ACCTCCAAGT
 4851 CTCAGAACAT TCTCTGCCT CACCTGGGC CACAGGTGT CGGCACCAC
 4901 CCCGACAAT TTTTGTGTT TTAGTAGATA TGGGGGTCTC GCTATGTTGC
 4951 CCAGGCTGGT CTAAACTCT TGGACTCAAG CGATCTTCCC ACCTCGGTAC
 5001 TAAAAAGTGC TGGGATTCCA GGTGTGAGCC ACCGTGCCCA, GCCTAGGTCC
 5051 TACTTTTATC TCCAATTTCAC AGATGAGTCC ATTTGAGAGA AGCTGACCC
 5101 CTTGCCCTGG GTCTCAAGGC TGGGGCGTGG CAGCACTTGG GTCCACGTTT
 5151 GTGCCCTTTC TGAATTCAG GACAACGTCA AAGATGGTCC TCAACCCAAAT
 5201 CCTCTGGGCT CCTCCAGTGG GGTAGTGGG TCCCTGGTGC ACACAGAAA
 5251 GCCTTGGGAGGCTGAATGG GGGTCCCCCG ACTCTGGCTT TCCCCCAGGT
 5301 ACATGTTCCCT GGTGAAATGG CAGGACCTGT CGGAGAAGGT GGTCTACCGG
 5351 CGCTTCACCG AGATCTACGA GTTCCATGTG AGTGTGGGG CGGAGGAGGG
 5401 ACAGGGACCC ACCGTTCCAG CTCCACCCCT TGGGAAGGAC CTTAGCCCAG
 5451 GTGATGGGGAA AACTGCAGAA CCCAGAACTC CCTCCCGACAC CACAGTTAAA
 5501 GGGGATTAT TTTTTTATAT AAATTTTTGT GACAGGGTCT TGCTCTGTCA
 5551 CCCAGGGTCT TGCTCTGTCA CCACTCTGAA CACCTCATGT TCTCTGATTA
 5601 CAGGCATGAG CCCCCACGGT CGGCCCTTTA GGTGGTTTG AGAGGTATTT
 5651 AGGTTTGCAAG TGCAAGGGCG CAATCATAGC TCACTGCAGC CTCACCTCT
 5701 GGGGCTCAAG CGATCTCCCT GCCTCAGCCT CCTGAGTAGC TGGGACTATA
 5751 GGTGCGCATC ACCATGTGTG GCTAATTTTT GTATTTTTA TAAAGATGGG
 5801 GATCTCACTA TGTGCCCCAG GCTGGTCTTG AACTCCAGAC CTCAGTGAT
 5851 CCTCTGCCT TGGCCTCCCA AAGCTAAGGG GGCATTTAAA GAAAAAAACAA
 5901 TTTTCCCCC TGAAACATTG AAGTAGTCTT ACTGAAAACA ATAAAACACA
 5951 GAAACACCAAG ATTCTCATTT TAAAGAAAA CAGACAGGAT CTCCCAAGAAC
 6001 CTTCTAGAA TGGAACCTT CTTGTCGTT TTGAAAAAA AAGCCAAGTT
 6051 CTAGCTCCA AATAAAATGCA CCTGCTGGT AACTCTCTC TTGTGGTTCT
 6101 CGTCCTATG TTAGTTATTT TCTTAAATTT TACATTGTAA CCTTTTTAAG
 6151 AATGAGTTAT CAGTTTTTTT ATATTTGCTT TTCTTTGAG ATGGGGTCTT
 6201 GCTCTGTCAAC CCAGGCTGGG GTGCAGTGGT GCAATCACGG CTCACGTGAG
 6251 CCTCAACCTC CAGGGCTGAA GCGATTCTCC CATCTCAGCC TCCCATGTG
 6301 AGATCACAGG TGTGACCAC CACACCTGGC TCTTTTCCCT GATTTGTTT
 6351 TTGAGAGAT GGGATTTCGC TATGTTGCCCT AGGCTGGTCT CTAACCTCTG
 6401 GACTCAAGTG ATCCTCCCGC CTCAGCTTCC CAAATTGCTA GGATTACAGG
 6451 TTTGAGCCCC TGCACCTGGT CAACCTGAGT TTTAAGAGGA TCCCTTTGGC

FIGURE 3B

6501 GACTGGATTG AGGACAGACA AGAGTGGACG GGGGACACAA GGAGGCCATT
6551 TTCGTTATCC AGGCCTGGTA GTGGCTAGGG CCAGGAGGGT GGGGTTGGTG
6601 GGAAGCAGTC AGATCCAAA GAGATTTGGG GATTGGAAGC AAAAGGATT
6651 GCTGGTGA CTCACATGGG AGGGAGAGAG GTCAGTGCT CTGTTAATCA
6701 AGGAATCCAG ATTGCCACCG AAATTTCTAG GCCCAGATA TTTAGGTAGT
6751 GTCTCACTCT GTCACCCAGG ATGGAGTGCA GTGGGCCAT CTGGCTCAC
6801 TGTAACCTCC GCCTCCAGG TTAAACAGAT TCTCCACCT CAGCCTCTG
6851 AGTAGCTGGG ATTACAGGCA TGTGCCACCA CTCCCGCTA ATTTTTGTAT
6901 TTTTAGTAGA GACGGGGTTT CACACGTTG GCCAGGCTGG TCTTGAAC
6951 CTGACCTCAA GTGATCCACC CACGACAGCC TCCAAAGTG CTGGGATTAC
7001 AGGGGTGAGC CACCATGCTC GGCTTTAG GTGGTTTGAG GAGGTATTTA
7051 GGTCACTTCC AATCTCGTC TTAACTCAAGT GTTGTAAACT ACAAAATATT
7101 CTTCACGTCT TCTTGTCTT TTAATGTITA GAAAACCTTA AAAGAAATGT
7151 TCCCTATTGA GGCAGGGGCG ATCAATCCAG AGAACAGGAT CATCCCCCAC
7201 CTCCCAGGTG AGCACGGGGC TGAGCCGCCT GTCAAGGGGT CATTGGCGGG
7251 GGCTCACCTG CCCTCCCGC ACCTCTCGGG CTTGACCTCA TGTTCCTCTGG
7301 TGCCAGCTCC CAAGTGGTTT GACGGGCAGC GGGCCGCCGA GAACCAC
7351 GGCACACTTA CGGAGTACTG CAGCACGTC ATGAGCTGC CCACCAAGAT
7401 CTCCCGTGT CCCCACCTCC TTGACTCTT CAAGGTGGC CCTGATGACC
7451 TCAAGCTCCC CACGGACAAAC CAGTGAGTGA ACTTTTCACT CTGCCAGGTG
7501 GGAGAGGGAA GGAGGGGTGG GACTTTCTGT GTTTTGAGA TGAGGAAACC
7551 AAGGCTCAGA GAGGGAAAGC CACCTTCCA GAGCACACA GCCAGAAAGA
7601 GGAGGCAAAT TCCACCTCCG GCCCCCTGTGA CCCCCCGCAAG CCTCCACCTT
7651 AATCTTCAC ACCTCAGGGC ACTGGGGAA GCACTCGGGG CTGGAGGTTC
7701 AAAGTCTGG GTCCCTCATCC TGACATTATG GCCACCTGGC CATGGGACCT
7751 GGAGCCAGTC ACCACTGCTC TCTGAATGCA GGTTCTCAT TTCTATAATG
7801 GGCAGTGAGG ATCAAGATGAA GCATTGGTG TCTTGGGAG CCCCCCAGAA
7851 GGATGTGGGG TTGATGCCCT TGCTAAGTGC TGAGCATGTC TGGGGTCTCC
7901 TGTACCCAGG ACCCTGTGTG GAAGGCACCT GAGAGGCTGA GGGAGCTCCA
7951 GGCAGGCTGG GGAAGTCCCC TTCTCCACTC CTCTCTGGTC ACTGAAGCTC
8001 GAACTGGGG ACGTGGGAC AGGACGTTAC CCCTTGTCAA GGCACCCAGG
8051 CTGCCAAGAC AGAGACAAGC AGCATTGCTC CGGCCAGCAC TTATTGACGC
8101 TTGAAGGTGT CCCCTGGCC AAGGAAGGGC AGTTATCATC AGCCCCGGGAG
8151 GCGGGGGAAAG GATGGACTCT GCAGTGGGT CCGCTCTCA TTGCTGTCTC
8201 TCTCAGGGCT CCAGAAGGAG GAAGAGGGCG GGCACAGTGG CTCACACCTA
8251 TAATCCCAGC ACTTTGGAAAG GTCGAGGTGG GCAGATCACC TGAGGTTGGG
8301 AGTTTGAGAC CAGCTGGCC AACATGGTA AACCCCATCT CTACCAAAAAA
8351 TATAAAAATT TAGTCAGGCA TGGTGGTGTG CGCTTGTAAAT CCCAGCTACT
8401 TGGGAGGCG AGCAGGAGA ATCGCTTGA CCGGGGAGGC AGAGGTTGCA
8451 GTGAGCTGAG ACTGGCCAC TGCACTCCAG CCTGGGTGAC AGAGCCAGAC
8501 TCTGCTTAAG AAAAAGAAA GAAAAGAAG AAGAAGATGG CCTGGGAGCC
8551 CGCAAGAGCA TTTCAGGC TTAGGGCCTC CTTTGGGTCT GCAGAAGGCT
8601 ATGCAGTGTC CTCTCTGT CCCTCCCTTG GGCTGCCGA GCAGATCCGC
8651 CGGCCCCCAT CACTCTGA AGCCCTTCT CAGCCAGTCC AGTTGCTGTC
8701 TTCTCTCCGC AGTGCCTTCC CGCTTTCCCG GGTCCCTCTT CTCTTGGGAA
8751 GTTCTCTGC AGGTCTACCC AGTGCTCTT CTTCTCCAT GGGAGCCAA
8801 GGGTCTCACC CAGACTGTT CTCCTCTCAGG ACAAAAAAGC CAGAGACATA
8851 TTGATGCCA AAAGATGGCA AGAGTACCGC GACAGGTGAG AGGACGGGGG
8901 GCAGCCGGCG GGGGGGACA CCCTGAGGAG ACCCAGAGTG TTCAGGGAAAT
8951 GGAGCAGGGG CTGGGAGCAG GCTGGGAGGG CTCACAGCTA CCCTGCTGAA
9001 GAATTGGGT TTTGGGCGG GTGCCGTTGC TCATGCCGT AATCCCAGCA
9051 GTTTGGGAGG CGGAGGAGG TGGATCACTT GAGGTAGGA GTTGAGACC
9101 AGCCTGGCA ACATGGAGAA ACCCTGCTC TACTAAAAT CCAAATTAGC
9151 CAGGGCTGGT GACAGGTGCC TGAGTCCA GCCACTTGGG AGGCTGAGGC
9201 AGGAGAATTG TTGAAACCCG GAAGACGGAG TTGAGTGA GCCGAGATCG
9251 TGCCACTGCA CTCCAGCTG GGCAGCAGAG CCAGACTCCA TCTCAAAAAA
9301 AAAAAAAAAGAAGAATT GGGTCTTGG AAGGTCCCTG GAGACTGAAA
9351 GGAGCCCTT GCAAGTGGCA GTGCAGAGAC CAGCCAGAC CCTTGCTACT
9401 GGCAGCCGGG GGAGTGGTT CGGCTGAATG AATGAACAGG TTTGGAGGG
9451 CAGGGTGGCC TTCAAGAGGC ATGCAGGGCT GTGGCAGTTT CTAATAC
9501 TTGCACAGTC ACTGCTAATA ACAATAATAA TAATAATACC TAACATTAAT
9551 GGAGTGCTTA CTCTGTGCCA GCCACTATTT TGTTTTGTG GTTTCACTG
9601 ACAGGGTCTC GCTCTGTGC CCAGGCCAGA GTGAAGTGGT GTGATCATAG
9651 CTCACACAG CCTCGACCTC CTGGGCTGAA GCGATCCCTC CACCTCAGCC
9701 TCCCAAGTAG CTGGGATTAC AGGTGTGTGC CACCATGTCC AGCTAATTT

FIGURE 3C

9751 TAATTTCTG ATAGAGATGG GGTCTCACTA CATTGCCAG GCTGGTCTTA
 9801 AGCTCTGGC CTCAGCAAC CCTCCTGCCT CAGCCTCCA AAGTGCTGAG
 9851 ATTATAGACA TGAGCCACTG TGCCCCGGCTT TTTCTCTTC TTATAAGGAC
 9901 ACGAGGCTG TTGGGTTAGG GCCCACTCTA CTGACCTCAT TTTAACTTAA
 9951 TTACCTCTG AAACGTACTT AAGAGTACCT TTCTCTTAAT ACACCCCACAC
 10001 TGTAAGGTAC TGGGTGGTTA GGACTTCAC ATATGAATT TGAGAAGGGC
 10051 GATGTCAGCC AATACCAAAC AGCATCAGCA CCTCCACGGT TGATGAAGG
 10101 GCTGGTCAGA AATGCACACT CAGGTCCAC AGTGACCTA CTGAACAGGA
 10151 TAGGCATTT AGAAAATCC CAGGTATTCTG GGTGACCTT AAAGTCTAGGA
 10201 AAAGGTCAGG CACTGTGGCT CATGCCGTGA ATCCCAGCAC TTGGGGAGGC
 10251 CGAGGCGGTT GAATCACCTG AGGTCAAGGAG TTGAGACCA GCCTGACCAA
 10301 TATCGTAAA CTCCATCTCT ACTAAAAATA CAAAATTAG CCAGGTGTGG
 10351 TGGGGGTGTC TTGTAGTCCC AGCTACTTGG GAGGCTGAGG CAGGTGAATT
 10401 ACTTGAAACCT GGGAGGTGGA GTTGCAATG AGCCAAGATT GCACCACTGC
 10451 ACTCCAGTGA CAGAGCGAGA CTCCATCTCA AAAAAAAA AAAAAAAAGT
 10501 TGGGAAAAGG CCAGGTGCAG TGGCTCCACG CCTGTAATCC CAACACTTTA
 10551 AGAGGCTGAG GTGGGAGAAT CCTTTGAGCC CAGGAGTTCG AGACCAAGCCT
 10601 GGGCATTGTC CAAGACCTT GTCTTACAA AAAATTAGCC GGGTGTGGT
 10651 GCATACGCT GTGGTCCCAG CTATTCGGGA GGCTGAGGCA GGGAGATTGC
 10701 TTGAGCCTAG GAGTCTAGGG CTGAGTGTAG CTGTGATCAC GTCACTGTAC
 10751 TCTAGCCTGG GCAACAGAGC AAGACTCTGT CTCCAAAAAA GAAAATAAAG
 10801 TTGGGAAAAGG CTCACTAATC TCATCAGATG AGAACAAAGA CATGTTTGAA
 10851 GTGTGAGGCC GAAGCCTGGA GAACGCTAT CGCCCAAGGA ATGGCAGGGCA
 10901 GCAGAGACTC AAGATGCCAG CGCCTTCTCT GGAGGCCCAG ATGGGCCCCCTG
 10951 CAATGCCAC TCACCCCTGCC CTCCCTCTTG CCCAGACAT CACCGGGCCCC
 11001 ATCATCCTGC AGACGTACCG CGCCATTGCC GACTACGAGA AGACCTCGGG
 11051 CTCCGAGATG GCTCTGTCA CGGGGGACGT GGTGAGGTC GTGGAGAAGA
 11101 GCGAGAGCGG TCAGACCTCC CACCTTACGG GGCTCTTCC CCTGGTGTCTC
 11151 AGGAACCCAC AGCCACAAGC CCCCTGCCAA GGCTCAGGCA GCCTGGCCCC
 11201 TGGGAGGACT CCAGCTCTGT TAGGGGCCCT AAATGTCCTC CCCACACTGT
 11251 GGGTCGCTT CTCTCTTAGT GTGCACCTG TGGTGGCTGT GGGCATCTGT
 11301 GCATGGCAGG CCGGGGCGGG GCATGTCG TGTTCTGTC TGATGGGTA
 11351 TGGGACCGTC TGTTCTATT GAAGTGGCT CAGAGCTGTG ATTCTGTGAG
 11401 CATGTGTGCA TGATGCGATG TGACCTCATT GTCCAGTGTG GTGAAGGGTGA
 11451 CATTCCAAA TCTGAGCATT GGACATCAGT GTGTCTGTG CCTGTGTGTC
 11501 TCACCATCCC TGATGGCTGC AGGGAGCCGC TGGGCCCCCTGC CCTCAGTC
 11551 CATTCCCGCA CCTCTGGCAC AGGTTGGTGG TTCTGTCAGA TGAAGCAAA
 11601 GCGAGGCTGG ATCCCAGCAT CCTTCCCTCGA GCCCCCTGGAC AGTCCCTGACG
 11651 AGACCGAAGA CCTTGAGGCC AACTATGAG GTGCCCCCTG CCTCCCGAGG
 11701 CTGTAGGGT GTGGGAGAAA GGGGCAAGCA GGCTCAGGG ATATTGAGTG
 11751 ACTGCTTTGG AGTCTGGCTG GTTGCTGGC TTGGCAGAAA AGTCAGGGCT
 11801 AAGATCTCAT CGCTCTGGC TTGGGGGGCC TGCGAGTTG TGATGCCCTT
 11851 GGTCTGGACA GGGAAACCAGG AGGAGGAGCA GACGACTCGG GAGAGTGGGA
 11901 GGCCAGTGGT GTCTGTGGAT ATGTGGCAG GTTCAGTGGG AAGCTGAAGG
 11951 ATGAGCAGAC CTTAGGCTCA GGAAGGAGGG CTGCCGGAA GTGGGGGCAT
 12001 CATCACTGAC CAGAAAGGGG AAAACTGGCAG TGCCAGGGCT GGATGGGGCC
 12051 TGCATTGAGC TTGAAAAAAA CTATAATAGA ATTGGTTACC ATTITTTTTT
 12101 ATTATTTATT TATTTTTTACTTTTGATAGAGTCT CACTCCCTTG
 12151 CTAAGGCTGG AGTCCGGTGG TGCTATCTCA GTCACACTGCA ACCTCTGCCT
 12201 CCCAGGATCA AGTGTATTCTC CAGCCTCAGC CTCCCCAGGT AGCTGGGATT
 12251 ACAACGATGC ACCACCATGC CTGGATAATT TTTGTATTT TAGTTGAGAC
 12301 GGGGTTTCAAC CAGGTTGGCC AGACTGGCT CGAACCTCTG ACCTCAGGTG
 12351 ATCTGCTGC CTGGGCCCTCC CAAAGTGTG GAATTACAGA TGTGAGGCCAC
 12401 TGTCCTGGC CTGGTTACCC ACATTTTAAATGAGTGTAT TTCACCCCTT
 12451 TATGTGGATT TACAGCTTGT TTTTTTTTTTTTGAGAC AAAGTCTGGC
 12501 TCTGTCAACC AGGCTGGAGT GCAGTAATGC AATCTCAGCT CACTGCAACC
 12551 TTAGGCTCCT GGGTTCAAGC AATTCTCTG CCTCAGGCCAC CTGAGTAGCC
 12601 TGGGGTTACA GGCGATGCC ACCACGCCAG GCTAATTTTGATTTTTA
 12651 GTAGAGATGG GGTTTCGCCA TGTTGGCCAG GCTGGTCTCG AACTCCTGAC
 12701 CTCAGGTGAT CCGCCCGCT TGGCCTCCCA AAGTGTAGG ATTACAGGTG
 12751 GGAACCACCT TGCCCAAGCT GTGGCTATCG TTAAACACT GGGAGGGCT
 12801 GCAGCCCCA GGCGACAGT TAGCTGCAGC TGAGCAGTTC CCAGTGCCAG
 12851 GTAGACGGAT GCTCCACCCA CCTACTCATG GCTGATCTCT TGTCACTGT
 12901 AAGTGTCTGG ACAGACCTTC ATCGTTATGG GATCTCTGGT CCCCAGAGTG
 12951 GGTGGCAATG AATGGGAGTG GACAAGCTCA CCTGGGTGTA GGGGGCAGAG

FIGURE 3D

13001 GGCGGAAGTC CAGAGTGTAC CCCCAGAGTG GGTGCCAGCA GGAGCTTGCC
 13051 GAGGATCTG GGATGGAGCA GGAGGGTGA GGGAGGAGAC CCAGAACAGG
 13101 GGGAACTGTG GGCCCTGGGT GGGTCTGGAG TGCTTGGAGG AAGCCCAGGC
 13151 GCAGAGAGGA GAAGATGGGA TGGGTGGCA GCCCCAGGCT GGCCGACCT
 13201 CACACTGTGC TCTGTGCCCT TGCCGTGGAC CAGGTGAGCC ATACGTGCGC
 13251 ATCAAAGGCT ACACGTGTG GGAGGGGAC GAGGTGTCCC TGCTCGAGGG
 13301 TGAAGCTGTT GAGGTAAATT ACAAGCTCT GGACGGCTGG TGGGTCATCA
 13351 GGTAGGAGGG CCCCTCTCCA TCCAGGAC CCATCTGAGT CAGCCCCAGC
 13401 CAGGACGGGG TGTAGGGGA TCTGGGGTGA CTGTCCCCTG GGACTCTGGG
 13451 TAAGGCACTG CCCCTCTCTG GGCTTAGTTT CCATCTCAGT AGCAGGGAGG
 13501 GATGAGCCA CCCTTGCTG TCTTGTGGG ATCCAATGTC CTTGTCCAAG
 13551 TGGGTGATT TCTCTTTGT GATTTAGGGT CTCTTCCCAA CCATCTTATA
 13601 TTATTCCCTC TCTGGCAACA TGGTGAACCTG TTGTATAAAT AATTACATTC
 13651 CTAGCTAGGC GCAATGGCCC AGGCTGTAA TCCCAGCACT TTGGGAGGCC
 13701 AGGACAGGAC GATCACGTGA GGTCAAGGAGT TCGAGACAC CCTGGCCAAC
 13751 ATGGCAAAAC CCTATCTCTA CTAAAAAACAC AAACATGAGC CGGGTGTGTT
 13801 GGTGGGAGCC TGTAAATCCCA GCTACTCGGG AGTCTGAGAC AAGAGAATCA
 13851 CTTCACCCG GGAGGCGGGAG GTTGCACTGA GCCAAGATCG CGCCATTGCA
 13901 CTCCAGCCTG GGCACAGGAGA GCGAAACTCC GTCTCAAAAAA AAAAAAAAAA
 13951 AAAAAAGATT ACCTTCTTTT TATCATTCTT TTATCTTTA AAGCTTCTT
 14001 GCAGTCAGGT GCAGTGTCTC ATGCCGTAA TCCCAACACT TTGGGAGCT
 14051 GAGGTGGGAG GATCACTAA GGCTACAAGT TCAAGACCAA CCTGGGCAAT
 14101 GTAGGGAGAC CTCGTCTCT ACAAAAAAAA TTAAAAAATA GCTGGATGTG
 14151 GTAGCACACA CCTGTAGCCC CAGCTACTCA GGAGGCTGAG GTGAAAGGAT
 14201 CACTGACCC CAGGAGTTGG AGGCAGCAGT GAGCTATGAC TGACCACCTG
 14251 CACCCAGCC TGGGTGATGG AGCAAGACCC TGTCTCAAA AAAAAAAAAA
 14301 AAAAAAGCT TCCATTGCAA TTCCCATCTG TTTATCCTCC AAATGAATGC
 14351 AGAAATACTA ATTATCTTTT TTCTGGTTCT GGGGAACACA GAATTCTAGC
 14401 GGCTTGTGGA GCCATTTCCC TGGAGCCATG GGGCCTCCCA GGTCCCTTCC
 14451 TGTGCTTCA TTTTTACGA ATTTTTCAT TTTTGAGAC AGGATCTTGC
 14501 TCTGACTCCC AGGCTGGAGC ACAATCATCG CTCACTCAAG CGATCCTCCC
 14551 ACCTCAGGCT CCCACGTAGC TGGGACTACA GGTGAGCACC ACCACATCTG
 14601 GCTAATGTTT TTAATTCTT TTGTAGGGT GGGGCTCTCAC TATGGTGCCA
 14651 AGACTAGTCT TAAACTCTG GCCTCAAGAG TTCCCTCTGC CTTGGCCTCC
 14701 CAAAGCACTG GGATTACAGG AATGAGCCTC CATGCTGGC CTTTGTCTGG
 14751 GTCTTCAGAG CCCTAGGTCA CAGGGCCAGC CTGGGCCCT GCCGCAAGCT
 14801 TATCTTAAAG CTGGGACAC AACATGCTATA CCTGCAAGCCGG GGGCCGGGGC
 14851 CAGAGGGCTT TGAGGCAGCA TTCTCAGCC TTTTAGACAC AACTCTGTT
 14901 AACCCCCATC CTGTGTCTCT GATAATCTTC TTGTGATCCT CCCACCAAGC
 14951 AAGAATTGGG TTATGTGA ACCTTGATT ATGCAAAGTT TTCTTTGTT
 15001 TTTTTTTCA CTCCAAAATA TAATATTGAG AATAGAAAGA AAGTCTTTTC
 15051 AACAAATGGT GCTGGAACAG ATGGATTTCCTC ATACTGGAAA AAAAAAAAAA
 15101 AGAGAAAAAA ACAAAACCTAG ACCCCCTCCCT CACACTGTAC ACATATGTTT
 15151 ACTTCAGATG GATCACAGGT TTATCCAGA GTAAAACCTG AACTAAAAAA
 15201 CCATTTGGGG CTGGACAGGG AGCTCACGCC TGTAATCTCA GCACTTTGGG
 15251 AGGCTGAGGC AGGTGGATCA TTGATGTCA GGAGTTTGAG ACCAGCCATG
 15301 ACCAATATGG TGAAATCTG TCTCTACTAA AAAAATACAA AATTAACCAA
 15351 GTGTGGTGGT GCATGCCGT AATCCCAGCT ACTTGGGAAG CTGAGACAGG
 15401 AGAATTGCTT GAACTTGGGA AGCAGAGGTT GCAATGAGTC GACATCATGC
 15451 CATTGCACTC CAGCCTAGGC AACAAGAGCA AAAACTCTGTC TTGGGGTTGG
 15501 GTGGGGAAA AGCATTGGGA AGAAAGCATA GAATTGGTG GCTTGGAGGT
 15551 AGGCAAAGGT TCGTAGGAGA CAGAAGGAG TTAACATAAA AGAAAAATTG
 15601 GCAAATATAA TCCGTGCCAGT GTCTTCTTT TTCTTTAATT TTTTGGGAG
 15651 GTAGAGATAG GGGTCTTGCT ATGTTACCCA GGCTGATCTC CAACTCTGG
 15701 CCTCAAGCGA TCTCTCCACC TAGATCCCTC AAAGTACTGG GATTACAGGC
 15751 GTGAGCGACC GTGCCCTGCC CATTCTGCC AATGCTTAT AGCAAATACC
 15801 TGTCCTCTGC GGTGACCTGG ATCTGCTAAC CTCCACCCCT GCCTAGACTG
 15851 TGGAGGATT GCTGGAAAGGG TCTCAGTTGC ACAGACAGG AACTGAGGC
 15901 CCACAGAGC AGGTGTCCGG TTGTTTGAA CCTCTCAGCC TGTGCTAAC
 15951 CCAATTGTT AGAGAGAGCC CTGAAACCTT CTCCCTGGG CGCCCCCAGG
 16001 TGACTGCCCT AGCCTCAAGG GCTGCCCTGT TTGCAAGGAAA GACGACGTCA
 16051 CAGGTACTT CCCGTCCATG TACCTGCAA AGTCAGGGCA AGACGTGTC
 16101 CAGGGCAAC GCCAGATCAA GCGGGGGGG CCGCCCCGCA GGTAGCGGG
 16151 GGTCCCCGGG GCTGGGGGGG GTCGAGCGGG GCGCACCAAGG GGTTCGCTCT
 16201 GTCTAGGCCA TAGCTGGCA GTGCCGGGC GGGGGCTCTC AGCCTGGCAG

FIGURE 3E

16251 GAGAGGCAGG ACCCTCACGG GGGAAAGGGG CTGGACGCCGC CTGGCCGCGG
 16301 TGTGGGGCTG GCACGGGGGC CGAAGGAAAG CGGCATGCC CGGGGGCTTT
 16351 GGGGATGGGC AGTCCAGGGG GGCTCCCCGG AGAGGGGGAC GACAGACCGA
 16401 AGGCTGGTGA GGGGCGTGG AAACCGCCCA GGCTCTGCTG CAGGGCAAGG
 16451 GTCTTGTCG TGACGGGGGC AGCCGCCTCT TGTCCCGCCG GGGTCGTGCA
 16501 GACTACCGGC CCCCTACTGC CCCCCCACTTC CTCGGACCAAG GGGTGCCCCAT
 16551 CTGAGTCCTT GGGGGCAGGG GCGCCCTCGG GCTTGTACGA CGCCCCGTCC
 16601 CGCTGGCCA GGTCTGCCAT CCGCAACCGG CACAGCATCC ACCAGCGGTC
 16651 GCGGAAGCGC CTCAGGCCAG ACGCTATCG CCGCAACAGC GTCCGTTTTC
 16701 TGCAAGCAGCG ACGCCGCCAG GCGGGCCCG GACCGCAGAG CCCCAGGGAGC
 16751 CCGCTCGGTG AGTGCAGCGG GAGAGGGCAG GAAGGGCAAG CCCTAGGGG
 16801 GGAGTCAGCG GGAGAGGGCG GGCCAGAGGC AGGGCCAGAG TAGCGGGGGCG
 16851 GGACCAGAGG GCGGAATCAG AGGGAGAGGC GGGGACTGGA GGCGGGGGCA
 16901 GAGGAGGAGC CAGCGCTAGG GGGCGGAGCG ATCCCTAAGA GGCGGAGTCA
 16951 GAGGAGGAGG CACAAGCGGG AGGCAGGCC AGAGCGCGGA GCAGGAGTTG
 17001 GAGACCGCGG CGGGGCGAGG CCAGAGAGCG CTGTGGGCGG GGCCAGTGTG
 17051 CGGGCGGGG CGTCTGACTC GGCCCCGCTC TCTGCCGCA GAGGAGGAGC
 17101 GGCAGACGCA GCGCTCTAAA CCGCAGCCGG CGGTGCCCCC GCGGCCGAGC
 17151 GCCGACCTCA TCTGAACCG CTGCAGCGAG AGCACCAAGC GGAAGCTGGC
 17201 GTCTGCCGTC TGAGGCTGGA GCGCAGTCCC CAGCTAGCGT CTCGGCCCTT
 17251 GCCGCCCCGT GCCTGTATAT ACGTGTTCTA TAGAGCCTGG CGTCTGGACG
 17301 CCGAGGGCAG CCCCGGAGCCC TGTCAGCGC GGCTCCGGG ACCCTCAATA
 17351 AATGTTGCTT GGAGTGGACC GAGGCTCTGC AGGAATGCA GGAGGGCCCG
 17401 GCTCCGCCCC AGGGTTATT TCTAAGTTGA GGACAGGGAG GTTGTGAGTT
 17451 CTGNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 17951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 18001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 18051 NNNNNNNNNN NNNNNNTAAA AATTAGCTGG GCGTGGTGGC ATGCATCCAC
 18101 AATCCAGCT ACTGGGGAGG CTGAGGCATG AGAACCGCTT GAAACGGGGG
 18151 GGCAGATGTT GCACTGAGCC GAGACGGCGC CACTGCACTC CAGCCTGGAC
 18201 TACAGAGCGA GACTCTATCT CAAAAAAA AAAAAAAA AAGTAACCTA
 18251 GGTGAGGGT GTCTCTGTT ATTCACTGAG ACCGTGCCCC GGTATGAGG
 18301 TTGTACCGA AAGCAAGTAT TCACTATGCA CACTATTAC CGCTCACCT
 18351 AGCATTGAAG CCAGCTGTG GCCTGAAAGC CTTGCTTTG AGGGCAGGTC
 18401 TTTCCCCAAA ATGCAGACAC GAAGGTGCAA AGTGAAGCTG CCAGTCTTGC
 18451 AAAAGATGTA ACTTGTCAAG AAGGCCACGA GTGGCAGGG AAGCTGTCCC
 18501 ACATTTGGCG AAGTGGCTAT GTGAGGACGG GGGAGGGGG TCCTTAGAG
 18551 ATAAGAGACA ATCATAAGGG GAGATATCG AGAAAATCGT AAGGGGAGCA
 18601 GATGGTTGTC AAGAGAATAG GCTGACCATC GAAGGACTGG CAGAAGTTT
 18651 CAGAAAACCA CTGGACGGCT GGGCACAGTG GCTTAGGCCT GTAATCCAG
 18701 CACTTTGGG GGTGACGCA GGTGAATCAC TTGAGGTCAG GAGTTCCAGA
 18751 CCAGCCTGGC CAACATGGTG AAACCCCATC TCTACAGAAA ATATAAAAAT
 18801 TAGCCAGGCG TGGTGGCACA AGCCTAGAAT CCCAGCTACT TGGGAGGCTG
 18851 AGG
 (SEQ ID NO: 3)

FEATURES:

Start: 2038
 Exon: 2038-2109
 Intron: 2110-4800
 Exon: 4801-4805
 Intron: 4806-5298
 Exon: 5299-5377
 Intron: 5378-7131
 Exon: 7132-7207
 Intron: 7208-7306

FIGURE 3F

Exon: 7307-7472
 Intron: 7473-8829
 Exon: 8830-8885
 Intron: 8886-10986
 Exon: 10987-11109
 Intron: 11110-11572
 Exon: 11573-11680
 Intron: 11681-13233
 Exon: 13234-13339
 Intron: 13340-16036
 Exon: 16037-16141
 Intron: 16142-16611
 Exon: 16612-16757
 Intron: 16758-17091
 Exon: 17092-17210
 Stop: 17211

SNPs:

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
139	T	C	Beyond ORF(5')			
262	T	C	Beyond ORF(5')			
43	A	G	Beyond ORF(5')			
344	G	A	Beyond ORF(5')			
721	A	G	Beyond ORF(5')			
1038	A	-	Beyond ORF(5')			
1468	A	G T	Beyond ORF(5')			
2758	G	T	Intron			
2890	G	A	Intron			
3304	G	A T	Intron			
3896	C	T	Intron			
3906	-	C G	Intron			
3907	-	T	Intron			
3911	-	T	Intron			
3932	-	A	Intron			
3934	-	T	Intron			
3949	C	T	Intron			
3994	A	G	Intron			
6272	C	T	Intron			
6427	T	C	Intron			
6694	T	C	Intron			
7741	C	T	Intron			
8294	A	G	Intron			
9313	A	- G	Intron			
10838	G	A	Intron			
11093	G	A	Exon	187	E	E
11195	T	G	Intron			
11213	G	A	Intron			
11263	C	G	Intron			
13707	G	A	Intron			
14629	G	A	Intron			
14698	T	A	Intron			
16095	C	T G	Exon	284	P	S A
16266	C	T	Intron			
16629	C	T	Exon	305	P	L
16642	C	T	Exon	309	H	H
18537	C	T	Beyond ORF(3')			
18589	G	A	Beyond ORF(3')			
18720	G	A	Beyond ORF(3')			
18782	C	T	Beyond ORF(3')			
18841	C	T	Beyond ORF(3')			

FIGURE 3G

Context:

DNA
Position

139	TACTAAAAATACAAAATTAGCCAGGCGTGGTGGCGCACACCTGTAATCCCAGCTACTTGG GAAGCTGAGGCAGGAGAATCGCTTGAACCTGGAAGGCAGAGGTTGCAGTGAGCCGAGATT GTGCCACTGCACTCCAGCCTGGGCAACAAGAGCGAAACTTCGCTTCAAACAAATAAAATTAA ACGCCAGCATGTCITGGCTTCATGCCAGACCTCAACCCCTCACCCCCAGGAGATCAG GTCCGGACCATGAGCTGACCC [T, C] GGACTCAGGCAAGGGTGAAGTTGGTGCAGCCCTGGCTGCTGGGAGGCACAGGCTGAGCA GGCTGCCCTGGGCTGAGGCCCCCACTCATGAACCTCATGACCTTGAATGAGCTCAAAG CTCTGGGCTCCCAAGGCTCTAGGGGAGTGGGAGAGAGAGGGCTCAGCCTGTCCTGGG ATGCTGCCCTCCTCACCTCTTGTCCAAATCCCCTCCTGGCAAAGCTGACAGTCTT AATATCACTTGGAGAAAATGAGTCAGCCCTAACAAATTCAATGA
43	TACTAAAAATACAAAATTAGCCAGGCGTGGTGGCGCACACCT [A, G] TAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATCGCTTGAACCTGGAAGGCAGAGGT TGCACTGAGCCGAGATTGTGCCACTGCACTCCAGCCTGGGCAACAAGAGCGAAACTTCG TTCAAACAAATAAAATTAAACGCCAGCATGTCITGGCTTCATGCCAGACCTCAACCC CACCCCCAGGAGATCAGGTCCGGACCATGAGCTGACCCCTGGACTCAGGCAAGGGTGA GGTGCAGCCCTGGCTGCTGGGAGGCACAGGCTGAGCAGGCTGCTGGGCTGAGGCC 344
344	TAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATCGCTTGAACCTGGAAGGCAGAGGT TGCACTGAGCCGAGATTGTGCCACTGCACTCCAGCCTGGGCAACAAGAGCGAAACTTCG TTCAAACAAATAAAATTAAACGCCAGCATGTCITGGCTTCATGCCAGACCTCAACCC CACCCCCAGGAGATCAGGTCCGGACCATGAGCTGACCCCTGGACTCAGGCAAGGGTGA GGTGCAGCCCTGGCTGCTGGGAGGCACAGGCTGAGCAGGCTGCTGGGCTGAGGCC [G, A] CCACTCATGAACCTCATGACCTTGAATGAGCTAAAAGCTTGGGCTCCCAAGGCTTAG GGGGAGTGGGAGAGAGAGGGCTCAGCCTGTCCTGGCATGTCGCCCCCTCCTCACCTCT TTGTCCAAATCCCCTCCTGGCAAAGCTGACAGTCTTAATATCACTCTGGAGAAAATG AGTCAGCCCTAACAAATTCAATGAACCAATTGCTTACTTGAGGATTGAACTCAAGT CTCACTAACAGTGTGCCATTTCGTCCAGTGTCACTGGCCCTCATCCACACACACC 721
721	AGGCCCTAGCCTGTCCCTGGCATGTCGCCCCCTCCACCTCTTGTCCAAATCCCCT TCCGGCAAAGCTGACAGTCTTAATATCACTCTGGAGAAAATGAGTCAGCCCTAAC CAATTCAATGAACCAATTGCTTACTTGAGGATTGGAACTCAAGTCTCACTAACAGTCT GCCATTTCTGCTCCAGCTGTCACTGCCCTCATCCACACACACCCAAGGATGAGCATCTA ACGCTTGATGACACTCCCATGCCCGTTCATTCACTCATTCACTCATTCACT [A, G] TTCAATTGACTCATTCACTCATTCACTCACTCATTCACTCACTCAGTGAATTTGAGTCA CGATCCAATATTATGGCTCTGTGTGAGGCACTAGATGGAGGGCTGGGGTAGAG CCCCCTGATAACCCGGTCACTGCCCTAGCTTCTGGACACACATTGTGGTAAGGGGAGAC TAAAAAAATTAAAGTCAGGCCAGGCACGGCTGCTCATGCCCTGAATCCCAGCACTTGGGAG GCCGAGGGCAGTGAATTACCTGAGGTCAGGAGTTCAAGACCAAGCCCTGGCAACATGGAGA 1038
1038	TTCATTCACTCACTCATTCACTCACTCACTCACTCACTCACTCACTCAGTGAATTTGAGTCA GGCCCTGTGTGCCAGGCACTAGATGGAGGGCTGGGGTAGAGCCCTGATAACCCGGT CATGCCCTAGCTTCTGGACACACATTGTGGTAAGGGGAGACTAAAAAAATTAAAGTC GGCCAGGGCAGGCTCATGCCCTGAATCCCAGCACTTGGGAGGCCAGGCGAGTGAAT TACCTGAGGTCAAGGAGTTCAAGACCAAGCCCTGGCAACATGGAGAAACCCAGTCTCAATT [A, -]

FIGURE 3H

	AAAAAAAAAAAAAAATTAGCCAGGTGTTGGCACATGCCGTGAATCCAGCTACTCAGG AGACTAACGCAAGAGAATTGCTTGAACCCAGGAGGCAGAGGTGCGGTGAGCCGAGATCG CGCCATTGCACTCCAGGCTGGAAACAAGAGCAGACTCCATCTCAAAAAAAAAAGT GGGAGGCAGGGCAGGGATCACTAGAGGCCAGTAGTTGAGACCATCTGGGAAACAT AGCAGGACCTGTCATAAAAAAAATTAAACCGGGCATGGTGGCACACA
1468	ACTCCAGCCTGGAAACAAGAGCAGACTCCATCTCAAAAAAAAGTGGGAGGCAG AGGCAGGAGGATCACTAGAGGCCAGTAGTTGAGACCATCTGGGCAACATAGCAGGACC CTGTCGTACAAAAAAATTAAAAAAATTAAACCGGGCATGGTGGCACACACCCGTAGTC CCAGCTACTCCAGGGCTGAGGCAGGAGATCGCTGGAGGCCAGGAGTTGAGGCTGAG TGAACGTGATCCCACCACTGCACTTAAGCCTGGATAACAAAGCAAGACCCGTCTCAA [A, G, T] AACAA TAGCA ATAATAAAGAAAAATTAAATGCAATTGCGATGCATCA GTGATAAGT GCTCTGAGAAAAGGAGGGCAGGAAGAGGTGAGAAAGGTATGAGGTTTGTATGCAATG TGAAGTTATCAAGGAAGGCTTCTCGGAAGAGGTGACATTGAGCAGAGAAATGGAGGAGA GTTATGGAGGGAAAGATGGTGAATGGGGGAACTGGTCAAGACCAGGAATATGGTCAAGG GGGGAAAGATGGTCAAGGGGACGCAAGCAATGCAAAGGCCCTGAGGCAGGAGCAGCTGA
2758	TCAGAGAGTGGCCCTCCAAGAGGGAGGCCTGGAAACTAAAGCTCTCTCCCCAGCT GCCCTGTAGTGTCAAGTTAGGTCTTATCCCTCCAGTAGGGTGACACCATGACAGGGGCCA ATAGATCCTCCCATCTGCTCCAGGAGGCTGGACAAATGCCCTGCTCAGACACACAAGT CCACTGGTCCCCATAATCCCCTAGGAAGGGCAGGGAGGAACATACATTAGGAAAATTGAAG CTTGTATGGAACATTAGTCTTATGTGCCAAGACCTTCTCTTTTGTTATTTTTGTT [G, T] TTTGAGACAGAGTCTTGTCTGTTGCCAGGCCAGAGTGCAGTGGCACGATCTCAGCTC ACTGCAACCTCCGCCCTCCAGGTTCAACTGGTCTCCCTGCCCTAGCCTCAGAGTAGTTG GGATTACAGGTGCCCAACCAACGCCCTGGTAATTTTGTATTTTTAGTAGAGACAGGGT TTCACCATGTTGCCAGACTGGTCTAAACTCCTGACCTCAAGTGTATCCACCCACCTGGG CCTCCCAAAGTGCTGGATTACAGGCATGAGCCACCGTGCCTGGCTGTTTTTGAAAT
2890	CATCTGCCCCAAGGAGGCTGGACAAATGCCGTGTCAGACACACAAAGTCCACTGGTCCC CTAATCCCATAGGAAGGCCAGGGAGGAACATTTAGGAATTGAGCTTGTATGAAAC ATTAGTCTTATGTGCCAAGACCTTCTCTTTTGTTATTTTTGTTGTTGAGACAG AGTCTTGTCTGTTGCCAGGCCAGAGTGCAGTGGCACGATCTCAGCTCACTGCAACCTC CGCCTCCAGGTTCAACTGGTCTCCCTGCCCTAGCCTCAGAGTAGTTGGATTACAGGT [G, A] CCCCACCAACGCCCTGGTAATTTTGTATTTTAGTAGAGACAGGGTTTACCATGTTG GCCAGACTGGTCTAAACTCCTGACCTCAAGTGTATCCACCCACCTGGCTCCAAAGTG CTGGGATTACAGGCATGAGGCCACCGTGCCTGGCTGTTTTTGAAATGAGGTCTGGAGT GCAGTGGTGGCATAGTTCACTGAGCCTCAACCTCCAGGCCAAGTGTATCCCTCTG CCTCAGCCCCCTGAGTAGCTGGGCTACAGGGCACACCAACCATGCCCTGGCTAGTTTAA
3304	CAAAGTGTGGATTACAGGCATGAGCCACCGTGCCTGGCTGTTTTTGAAATGAGGT CTGGAGTGCAGTGGTGGCATAGTTCACTGCAAGCCTCAACCTCCAGGCCAAGTGT CCTCCCTGCCCTAGCCCTTGAGTAGCTGGGCTACAGGCCACACCACATGCCCTGGCTA GTTTTAAATTTTGTGGAGATGAGGTTCAGTATGTTGTCAGGCTAATCTGAACCTC CTCGGCTTAAGCAACCTCTGGTCTAGCCTCCACAGTGTAGGATTACAAGCGTGAGC [G, A, T] ACCGTGCCTAGTCACCTTCTCTTTGTAACCTTCAGTTGAAATTCTAAATT ACAGAAAGGCTACTGGGTGCAAAACGGTACCGACTCCAAATAGTCCTTCACTCACCT TCATCCACACCTCTCTGGGGATATTCTGAATTATTGAGAGTGAGTTGAAGACG TGTTCCTTACCTCTAAATACTAGTTGTTGGCTTAAATCAAGGCATTCTTA CATATAACACACACGTGTCAAATCAGGAATTACATGGACAAAACACCAATTATCA
3896	CATTATCCACCCACAGACTTACTGAGGTCCCCGATTATCTGCTTGTCTCTGCAGT GAAAATTTTCAGGTCTAGGATCCAGTCAGGATCAATGTCATAGCCTTAAACCTCT TTAATCTGGATCAGTCTTTCTTTCTTTCTTTCTGGACACGGAAATCTCACTC TGTGCCAGACTGGAGTGCAGTGGTGCACATTGCCCTCTGCCCTGG TTCAAGAGATTCTCTGCCCTAGCCTCTGAGTAGCTGGGAATACAGGTGCGGCCACCA [C, T] CCCCAGCTGTTTGAGAGACAGGGTTTGCCTAGATTCTGGATCAGTCCTTTT TTTTTTTATGAGATGGAGTCTTACTCTGTCAACCCAGGCTGGAGTGCAATGGCACAATCT CCACTCACTGCATCCTCCGCCCTCCAGGTTCAAGCAATTCTGTCCTGCCCTCCAG TAGCTGGATTACAGGCATGCGCCACCATGCCGGTACTTTTGTTATTTTAGTAGAGA CAGGGTTTACCATGTTAGCCAGGTGATCTGAACTCTGACGTCAGGTGATCTGCCCG

FIGURE 3I

3906 CCACAGACTTTACTGAGGTTTCCCCGATTATCCTGCTTGTCCCTGCAGTGAAAACCTTT
 TTCAAGGTCTAGGATCCAGTCAGGATCAATGTCATAGCCTTAACCTCTTAATCTGGAT
 TCAGTCCTTTCTCTTTCTTTCTTTGGACACGGAATCTCACTCTGTCGCCAGA
 CTGGAGTGAGTGGTGCATCTCGGCTCATTGCAACCTCTGCCCTCTGGGTTCAAGAGAT
 TCTCCCTGCCCTAGCCTCTGAGTAGCTGGAAATACAGGTGCGGCCACCACGCCAGCTC
 [-, C, G]
 TTTTTGGTAGAGACAGGGTTTGCCTTGTGATTCTGGATCAGTCCTTTTTTTTTAT
 GAGATGGAGTCTTACTCTGTCAACCCAGGCTGGAGTGCAATGGCACAATCTCACTCACTG
 CATCCCTCCGCCCTCCAGGTTCAAGCAATTCTGTCGCCCTAGCCTCCCGAGTAGCTGGGAT
 TACAGGCATGCCACCATGCCGGTACTTTGTATTTAGTAGAGACAGGGTTCA
 CCTATGTTAGCCAGGCTGATCTGAACCTCTGACGTAGGTGATCTGCCGCCCTGACCTC

 3907 CACAGACTTTACTGAGGTTTCCCCGATTATCCTGCTTGTCCCTGCAGTGAAAACCTTT
 TCAGGTCTAGGATCCAGTCAGGATCAATGTCATAGCCTTAACCTCTTAATCTGGAT
 CAGTCCTTTCTCTTTCTTTCTTTGGACACGGAATCTCACTCTGTCGCCAGAC
 TGGAGTGAGTGGTGCATCTCGGCTCATTGCAACCTCTGCCCTCTGGGTTCAAGAGATT
 CTCCCTGCCCTAGCCTCTGAGTAGCTGGAAATACAGGTGCGGCCACCACGCCAGCTC
 [-, T]
 TTTTTGGTAGAGACAGGGTTTGCCTTGTGATTCTGGATCAGTCCTTTTTTTATG
 AGATGGAGTCTTACTCTGTCAACCCAGGCTGGAGTGCAATGGCACAATCTCACTCACTG
 ATCTCCGCCCTCCAGGTTCAAGCAATTCTGTCGCCCTAGCCTCCCGAGTAGCTGGGATT
 ACAGGCATGCCACCATGCCGGTACTTTGTATTTAGTAGAGACAGGGTTCAC
 CATGTTAGCCAGGCTGATCTGAACCTCTGACGTAGGTGATCTGCCGCCCTGACCTCC

 3911 GACTTTACTGAGGTTTCCCCGATTATCCTGCTTGTCCCTGCAGTGAAAACCTTTTCAG
 GTCTAGGATCCAGTCAGGATCAATGTCATAGCCTTAACCTCTTAATCTGGATCAGT
 CTCTTTCT
 GTGAGTGGTGCATCTCGGCTCATTGCAACCTCTGCCCTCTGGGTTCAAGAGATTCTCC
 TGCCCTAGCCTCTGAGTAGCTGGAAATACAGGTGCGGCCACCACGCCAGCTCGTTT
 [-, T]
 GGTAGAGACAGGGTTTGCCTTGTGATTCTGGATCAGTCCTTTTTTTATGAGAT
 GGAGTCTTACTCTGTCAACCCAGGCTGGAGTGCAATGGCACAATCTCACTCACTGCACTC
 TCCGCCCTCCAGGTTCAAGCAATTCTGTCGCCCTAGCCTCCCGAGTAGCTGGGATTACAG
 GCATGCGCCACCATGCCGGTACTTTGTATTTAGTAGAGACAGGGTTTACCATG
 TTAGCCAGGCTGATCTGAACCTCTGACGTAGGTGATCTGCCGCCCTGACCTCCAAA

 3932 ATTATCCTGCTTGTCCCTGCAGTGAAAACCTTTTCAGGTCTAGGATCCAGTCAGGAT
 CAATGTCATAGCCTTAACCTCTTAATCTGGATCAGTCCTTTCTCTCTCTCTCT
 TTTTTGGACACGGAATCTCACTCTGTCGCCAGACTGGAGTGAGTGGTGCATCTCGG
 TCATTGCAACCTCTGCCCTCTGGGTTCAAGAGATTCTCTGCCCTAGCCTCTGAGTAGC
 TGGGAATACAGGTGCGGCCACCACGCCAGCTCGTTTTGGTAGAGACAGGGTTTGCC
 [-, A]
 TTGATCTGGATCAGTCCTTTTTTTATGAGATGGAGTCTTACTCTGTACCCA
 GGCTGGAGTGCATGGCACAATCTCACTCACTGCACTCTGCCCTCCAGGTTCAAGCA
 ATTCTGTCGCCCTAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCATGCCGG
 TACTTTGTATTTAGTAGAGACAGGGTTTACCATGTTAGCCAGGCTGATCTCGAAC
 TCCTGACGTAGGTGATCTGCCGCCCTGACCTCCAAAAGTGTGGGATTACAGGCGTGA

 3934 TATCCTGCTTGTCCCTGCAGTGAAAACCTTTTCAGGTCTAGGATCCAGTCAGGATCA
 ATGTCATAGCCTTAACCTCTTAATCTGGATCAGTCCTTTCTCTCTCTCTCT
 TTTTGACACGGAATCTCACTCTGTCGCCAGACTGGAGTGAGTGGTGCATCTGGCTC
 ATTGCAACCTCTGCCCTCTGGGTTCAAGAGATTCTCTGCCCTAGCCTCTGAGTAGCT
 GGAATACAGGTGCGGCCACCACGCCAGCTCGTTTTGGTAGAGACAGGGTTTGCCAT
 [-, T]
 GATTCTGGATCAGTCCTTTTTTTATGAGATGGAGTCTTACTCTGTACCCAGG
 CTGGAGTGCAATGGCACAATCTCACTCACTGCACTCTGCCCTCCAGGTTCAAGCAAT
 TCTGTCGCCCTAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCATGCCGGCTA
 CTGGTAGAGACAGGGTTTACCATGTTAGCCAGGCTGATCTCGAACCTC
 CTGACGTAGGTGATCTGCCGCCCTGACCTCCAAAAGTGTGGGATTACAGGCGTGA

 3949 CTGCACTGAAAACCTTTTCAGGTCTAGGATCCAGTCAGGATCAATGTCATAGCCTTA
 ACCTCTTTAACTCTGGATCAGTCCTTTCTCTCTCTCTCTCTCTCTCTCTCT
 CTCACTCTGTCGCCAGACTGGAGTGAGTGGTGCATCTGGCTCATTGCAACCTCTGCC
 TCCCTGGGTTCAAGAGATTCTCTGTCGCCCTAGCCTCTGAGTAGCTGGGAAATACAGGTGCC

FIGURE 3J

GCCACCACGCCAGCTGTTTGGTAGAGACAGGGTTGCCATTGATTCTGGATCAGT
 [C,T]
 TTTTTTTTTTTTATGAGATGGAGTCTACTCTGTCACCCAGGCTGGAGTGCAATGGC
 ACAATCTCCACTCACTGCATCCTCCCTCCAGGTTCAAGCAATTCTCGTCCTCAGCC
 TCCCAGTAGCTGGGATTACAGGCATGCGCCACCATGCCCGGTACTTTTGTTA
 GTAGAGACAGGGTTTACCCATGTTAGCCAGGCTATCTGAACCTCTGACGTAGGTGAT
 CTGCCCGCCTGACCTCCAAAGTGCTGGGATTACAGGCAGTGGCCACCGTGCCAGCGGA

3994 ATGTATAGCCCTTAACCTCTTAACTGATCAGTCTTTTCTTTCTTTCTTT
 TTTGGACACCGGAATCTCACTCTGTGCAGCAGACTGGAGTGCACTGGTGAACTCTGGCTC
 ATTGCAACCTCTGCCCTCTGGGTTCAAGAGATTCTCCGTGCTCAGGCTCTGAGTAGCTG
 GGAATACAGGTGCGCGCCACACGCCAGCTCGTTGGTAGAGACAGGGTTTGCCAT
 TGATTCTGGATCAGTCTTTTCTTTTATGAGATGGAGTCTACTCTGTACCCAG
 [A,G]
 CTGGAGTGAATGGCACAATCTCACTCACTGCATCCTCCGCTCCAGGTTCAAGCAAT
 TCTCGGCCCTCAGCCTCCGAGTAGCTGGGATTACAGGCATGCGCCACCATGCCCGCTA
 CTTTTGTTAATTTAGTAGAGACAGGGTTTACCATGTTAGCCAGGCTATCTGAACCTC
 CTGACGTCAAGGTGATGCCCCCTGACCTCCAAAGTGCTGGGATTACAGGCAGTGAAC
 CACCGTGCCAGGGATTCTGGATCGGTAACTCAGTCTTTGTCCTTGCAACTTGATG

6272 AAAGTAAAACAGACAGGGATCTCCAGAACCTCTAGAATGAAACCATTCTGCGCTT
 TGAAAAAACAAAGCCAAGTTCTAGATCCAAATAATGACCTGCTGGTGAAACATTCTCCT
 TGTGGTTCTGTCCTATGTTAGTTATTTCTAAATTACATTGACCTTTAAGA
 ATGAGTTATCAGTTTTTATATTGCTTTCTTGAGATGGGTCTGCTCTGTCACC
 CAGGCTGGGTGCACTGGTCAACCTGAGTTAAGAGGATCCCTTGCGACTGGATTGAG
 [C,T]
 GATTCTCCCATCTCAGCCTCCATGTTAGAGATCACAGGTGTGCAACCACACCTGGCTC
 CTTTCTGATTGTTAGAGATGGGATTCTGCTATGTTGCCAGGCTGGCTCT
 AACTCCTGGACTCAAGTGATCCTCCGCTCAGCTCCAAATTGCTAGGATTACAGGTT
 TGAGCCCTGACCTGGTCAACCTGAGTTAAGAGGATCCCTTGCGACTGGATTGAG
 GACAGACAAGAGTGGACGGGGGACACAAGGAGGCCATTTCGTTATCCAGGCTGGTAGT

6427 ATTTTACATTGACCTTTAAGAATGAGTTATCAGTTTTTATATTGCTTTCTTT
 TGAGATGGGTCTGCTCTGTCACCCAGGCTGGGTGCACTGGTCAATCACGGCTCACT
 GCAGCCTCAACCTCCAGGGCTGAAGGATTCTCCATCTCAGCCTCCATGTTGAGATCA
 CAGGTGTGACCAACACCTGGCTCTTCTGTTGTTGTTGAGATGGGATT
 TCGCTATGTTGCCAGGTGGTCTAACTCCTGACTCAAGTGATCCTCCGCTCAGC
 [T,C]
 TCCCCAATTGCTAGGATTACAGGTTTGAGCCCTGACCTGGTCAACCTGAGTTAAGA
 GGATCCCTTGGCACTGGATTGAGGACAGACAAGTGAGCCAGGGGACACAAGGAGGCC
 ATTTCGTTATCCAGGCCGGTAGGGCTAGGGCAGGAGGGGGTTGGTGGGAAGCA
 GTCAAGATCCAAAGAGATTGGGGATTGGAAGCAAAGGATTGCTGGTACTTGACAT
 GGGAGGGAGAGAGGTCACTGCTCTGTTAATCAAGGAATCCAGATTGCCACCGAAATTTC

6694 ACTCCTGGACTCAAGTGATCCTCCCGCTCAGCTCCAAATTGCTAGGATTACAGGTT
 GAGCCCTGACCTGGTCAACCTGAGTTAAGAGGATCCCTGGCACTGGATTGAGG
 ACAGACAAGAGTGGACGGGGGACACAAGGAGGCCATTTCGTTATCCAGGCTGGTAGTG
 GCTAGGGCCAGGAGGGTGGGTGGTGGAGCAGTCAGATCCAAAGAGATTGGGGAT
 TGGAGAAAAGGATTGCTGGTACCTGACCTGACATGGAGGGAGAGAGGTCACTGCTCTG
 [T,C]
 TAATCAAGGAATCCAGATTGCCACCGAAATTCTAGGCCCAGAATTAGGTAGTGTCT
 CACTCTGTCACCCAGGATGGAGTGCACTGGCCATCTCGGCTCACTGTAACCTCCGCT
 CCCAGGTTAACGATTCTCCACCTCAGCCTCTGAGTAGCTGGATTACAGGCATGTG
 CCACCACTCCGGTAATTGTTAGTAGAGACGGGTTTACCAAGTGGCCA
 GGCTGGTCTGAACTCCTGACCTCAAGTGATCCACCCACGACAGCCTCCAAAGTGCTGG

7741 CCTGATGACCTCAAGCTCCCCACGGACAACCAAGTGAGTGAACTTTTACCCCTGCCAGGTG
 GGAGAGGGAAAGGAGGGGGGGACTTCTGTTGCACTGGAGTAGGAAACCAAGGCTCAGA
 GAGGGAAAGCCACCTTCCCAGAGCCACACAGGCCAGAAAAGAGGAGGCCAAATTCCACCTCCG
 GCCCCCTGTAACCCCGCCAAGCCTCCACCTTAATCTTACACCTCAGGGCACTGGGGAA
 GCACTCGGGCTGGAGGGTCAAAGTCTGGGTCTCATCTGACATTATGCCACCTGGC
 [C,T]
 ATGGGACCTGGAGGCCAGTCACCACTGCTCTGTAATGCAAGGTTCTCCATTCTATAATGG
 GCAGTGAGGATCAGATGAAGCATTGGGTGCTGCGGAGCCCCCAGAAGGATGTGGGGT
 TGATGCCCTGCTGTAAGTGCTGAGCATGTCCTGGGTCTCCTGTAACCCAGGACCTGTGTGG

FIGURE 3K

	AAGGCACCTGAGAGGCTGAGGGAGCTCCAGGCAGGCTGGGAAGTCCCCTCTCCACTCC TCTCTGGTCACTGAAGCTCGAAGTGAGGAGCATGAGGACAGGACGTACCCCTTGTCAG
8294	GAAGCTCGAACGTTGGGAGCATGAGGACAGGACGTACCCCTTGTCAGGCACCCAGGCTG CCAAGACAGAGACAAGCAGCATTGCTCCGGCCAGCACTTATTGACGCTTGAAAGGTGTCCC CTGGGCCAAGGAAGGGCAGTTATCATCAGCCGGAGGCGGGGAAGGATGGACTCTGCA GTGGGGTCCGCTCTCATTCGCTCTCAGGGCTCCAGAAGGAGGAAGGGCCGGC ACAGTGGCTCACACCTATAATCCCAGCACTTGGAAAGGTGAGGTGGCAGATCACCTGA [A, G] GTTGGGAGTTGAGACCAGCCTGGCCAACATGGTAAACCCCATCTTACCAAAAATATA AAAATTTAGTCAGGCATGGTGGTGTGCGCTGTAAATCCCAGCTACTTGGGAGGCCAGGC AGGAGAAATCGTGAACCCGGAGGGCAGAGGTGAGCTGAGACTGCCACTGCA CTCCAGCCTGGGTGACAGAGCAGACTCTGCTAAGAAAAAAAAAGAAAAGAAAAGAAA AGATGGCCTGGAGGCCAGCAAGAGCATTTCAGGCTTAGGGCATCCTTGGGTCTGAG
9313	TGGGCCGGGTGGGTTGCTCATGCCGTAAATCCCAGCAGTTGGGAGGCCAGGCAGGTG GATCACTTGAGGTCAGGAGTTGAGACCAGCCTGGCCAACATGGAGAAACCTGTCTTA CTAAAATCCAATTAGCCAGGCGTGGTACAGGTGCTGTAGTCCCAGCACTTGGGAG GCTGAGGCAGGAGAAATTGCTTGAACCCGGAAGACGGAGTTGAGTGCAGTGAGCCGAGATCGT CCACTGCACCTCCAGCCTGGGAGCCAGACTCCATCTCAAAAAAAA [A, -, G] AAGAATTGGGTTTTGAAAGGTCCCTGGAGACTGAAAGGAGCCCTTGCAAGGTGGCAGTG CAGAGACCAGCCAGACCCCTGCTACTGGCAGCCGGGAGTGTTGGGCTGAATGAAT GAACAGGTTTGAGGGCAGCGTGGCTTCAGAGGAGTGCAGGGCTGTGGCAGTTCTA ATACTTATTGACAGTCAGTCAATAACAATAATAACCTAACATTAAATGGA GTGCTTACTCTGTCAGCCACTATTGTTTGTGTTTCAAGTGCAGGGCTCGCT
10838	CTGTACTCTAGCCTGGGCAACAGAGCAAGACTCTGTCCTCCAAAAAAGAAAATAAGTTGG GAAAGGCTACTAACTTCATCAGATGAGAACAA [G, A] GACATGTTGAAGTGTGAGGCCAGCCTGGAGAACGCTATGCCCCAGAAAATGCAGGG CAGCAGAGACTCAAGATGCCAGCGCTGTCTG
11093	AAATAAAGTGGAAAGGCTACTAACTTCATCAGATGAGAACAAAGACATGTTGAAGT GTGAGGCCAGCCTGGAGAACGCTATGCCCCAGAAAATGCAGGGCAGCAGAGACTCAA GATGCCAGGCCCTGTTCTGGAGGCCAGATGGCCCTGCAATGCCACTACCCCTGCCCT CCCTCTGCCCAAGACATCACCGGCCCATCATCTGCAAGACGTACCGGCCATTGCCGA CTACGAGAACACTCGGGCTCCGAGATGGCTGTCCACGGGGAGTGGTGGAGGTG [G, A] GAGAAGAGGGAGAGCCGTAGACCTCCACCTTACGGGGCTCTTCCCCTGGTCAGG AACCCACAGCCACAAGCCCCCTGCCAAGGCTCAGGCAGCCTGGCCCTGGAGGACTCCA GCTCTGTTAGGGCCCTAAATGTCCTCCCCACACTGTGGGTCGCCCTCTCTTACTGTG CACCCCTGGTGGCTGGGCACTGTGCAATGGCAGGCCGGGGCATGTCTGCGTG TTCTGTCGGATGGGTATGGGACCGTCTGTCATTATGAAGTGGCTAGAGCTGTGATT
11195	AGGGCAGCAGAGACTCAAGATGCCAGCGCCTGTTCTGGAGGCCAGATGCCCTGCAAT GCCCACTCACCCCTGCCCTCCCTCTGCCCAAGACATCACCGGCCCATCATCTGCAAG GTACCGGCCATTGCCACTACGAGAACACTCGGGCTCCGAGATGGCTGTCCACGGG GGACGTGGGGAGGCTGTGAGAACAGGGAGAGGGCTAGACCTCCACCTTACGGGG CTTCCCTGGTGTCAAGAACCCACAGCCACAAGCCCCCTGCCAAGGCTCAGGCAGCCT [T, G] GCCCTGGGAGGACTCCAGCTGTAGGGCCCTAAATGTCCTCCCCACACTGTGGGTC GCCCTCTCTTAGTGTGCAACCTGTGGGGCTGGGCACTGTGCAATGGCAGGCCGG GGGGGAGTGTCTGCTGTTCTGTCGGATGGGTATGGGACCGTGTGTCATTATGAAGT GGGCTCAGAGCTGTGATTCTGTCAGACATGTGCACTGCATGCATGTGACCTCATGTCCA GTGTGGTGAAGGTGACATTTCCAAATCTGAGCATTGGACATCAGTGTGTCTGTG CTG
11213	GATGCCAGCGCCTGTTGGAGGCCAGATGGGCCCTGCAATGCCACTCACCTGCCCT CCCTCTGCCCAAGACATCACCGGCCCATCATCTGCAAGACGTACCGGCCATTGCCA CTACGAGAACACTCGGGCTCCGAGATGGCTGTCCACGGGGAGTGGTGGAGGTG GGAGAACAGGGAGAGGGCTAGACCTCCACCTTACGGGGCTCTTCCCCTGGTGTCA GAACCCACAGCCACAAGCCCCCTGCCAAGGCTCAGGCAGCCTGGCCCTGGAGGACTCC [G, A] GCTCTGTTAGGGCCCTAAATGTCCTCCCCACACTGTGGGTCGCCCTCTCTTACTGTG CACCCCTGGTGGCTGGCATGTGCAATGGCAGGCCGGGGCATGTCTGCGTG

FIGURE 3L

	TTCTGTCTGGATGGGTATGGGACCGTCTGTCATTATGAAGTGGGCTCAGAGCTGTGATT CTGTGAGCATGTGTCATGCATGCACTGACCTCATTGTCAGTGTGGTGAAGGTGACAT TTCCAAATCTGAGCATTGGACATCAGTGTCTGTGTCCTGTGTCCTCACCATCCCTGA
11263	ACCCCTGCCCTCCCTCTTGCCCCAGACATCACCGGCCCCATCATCCTGCAGACGTACCGCG CCATTGCCACTACGAGAAGACCTCGGCTCCGAGATGGCTCTGTCACGGGGACGTGG TGAGGCTGAGAAGAGCGAGGGTCAAGACCTCCACCTTACGGGGCTCTTCCC TGGTCTCAGGAACCCACAGCCAACAGCCCCCTGCCAAGGCTCAGGCAGCTGGCCCC GGAGGACTCCAGCTCTGTTAGGGCCCTAAATGTCCCTCCCCACACTGTGGTCGCCCTCT [C,G] TCTTAGTGTGACCCCTGTGGTGGCTGTGGGATCTGTGCACTGGCAGGCCGGGGGG TGTCTGCGTGTCTGTGAGGATGGGATGGGCTGTGTCATTATGAAGTGGCTAG AGCTGTGATTCTGTGAGCATGTGTGCACTGCACTGTGACCTCATTGTCAGTGTGG AAGGTGACATTTCCAAATCTGAGCATTGGACATCAGTGTGTCGTGTCCTGTGTC CCATCCCTGATGGCTGCAGGGAGCCCTGGGCCCCCTGCCCTGAGTCACATCCCGACCT
13707	GGGGTGTAGGGATCTGGGTGACTTGTCCCTGGACTCTGGTAAGCCACTGCCCTC TCTGGGCTTAGTTCATCTCAGTACAGGGAGGATGAGCCACCCCTGCCCTGCTTGT GGGGATCCAATGTCCTTGTCCAAGTGGGTGATTCTCTTGTGATTAGGGCTCTTC CCAACCATCTATTATTATTCTCTGCAACATGGTGAAGTGTGATAAAATAATTAC ATTCTCTAGCTAGGGCGAACATGGCCAGGCCCTGTAATCCCAGCACTTGGGAGCCCAGGACA [G,A] GACCATCACGTGAGGTCAAGGAGTTCAGAGACCAACCTGGCAACATGGCAAACCCCTATCT CTACTAAAAACACAAACATGAGCCGGGTGTTGGTGGAGGCTGTAATCCCAGCTACTC GGGAGTCTGAGACAAGAGAACATCTTCAACCCGGGAGGGCGAGGTTGCACTGAGCCAAGA TCGCCCTTGCACCTCAGGCTGGCAACGAGAGGAAACTCCGTCTCAAAAAAAAAAAA AAAAAAAAAGATTACTTTCTTTATCATTCTTATCTTTAAAGCTTCTTGAGTC
14629	TGTTTATCCTCCAAATGAATGCAGAAATACTAATTATCTTTCTGGTTCTGGGAACA CAGAATTCTAGCGGCTTGTGGAGCCATTCCCTGGAGCCATGGGGCTCCACGGCTT CCTGTGTCCTCATTTTACGAATTCTTCTGAGACAGGATCTGCTCTGACTC CCAGGCTGGAGCACAATCATCGCTCACTCAAGCGATCCTCCACCTCAGGCTCCACGTA GCTGGGACTACAGGTGAGCACCACATCTGGCTAATGTTTAATTTTTGAGGG [G,A] TGGGGTCTCACTATGGTCCAAGACTAGTCTAAACTCCTGGCTCAAGAGTCTCTCTG CTTGGCCTCCAAAGCACTGGGATTACAGGAATGAGCCTCCATGCTGGGCTTGTGG CGTCTCAGAGCCCTAGGTACAGGGCAGCCTGGCCCTGCCGAAGCTTATCTTAA GCTGGGACCAACACATGCATACCTGCAAGCCGGGCCCCGGGCAAGAGGGCTTGAGGCAGC ATTTCTCAGGCTTTAGACACACACTCTGTTAACCCCATCTGTGTCCTGATAATCTT
14698	TTCTTTTACGAATTCTTCTTGTGAGACAGGATCTGCTCTGACTCCAGGCTG GAGCACAAATCATCGCTCACTCAAGCGATCCTCCACCTCAGGCTCCACGTAAGCTGGAC TACAGGTGAGCACCACACATCTGGCTAATGTTTAATTCTTGTAGGGTGGGTC TCACTATGGTCCAAGACTAGTCTAAACTCCTGGCTCAAGAGTCTCTGCTGGC [T,A] CCAAAGCACTGGGATTACAGGAATGAGCCTCCATGCTGGGCTTGTGCTGGCTTCAAG AGCCCTAGGTACAGGGCAGCCTGGCCCTGCCGAAGCTTATCTTAAAGCTGGAC ACAACATGCATACCTGCAAGCCGGGCCCCGGGCAAGAGGGCTTGAGGCAGCATTCTCAG CTTGTAGACACACACTCTGTTAACCCCATCTGTGTCCTGATAATCTCTGTGATC C
16095	AATACCTGCCCCCTGCGGTGACCTGGATCTGTAACCTCCACCCCTGCCAGTGTGA AGGATTGCTGAAGGGTCTCAGTTGACAGACAGGAAACTGAGGCCACAGAGGCAGGT GTCCGGTTTGTGCAACCTCTCAGGCTGTCAACCCCAATTGTTCAAGAGAGGCCCTGA AACCCCTCTCTCTGGGCCCTGCCGAAGCTTACCTGCCCCAGGCTCAAGGGCTGCCCTGTTGC AGGAAAGACGACGTCAACAGGCTACTTCCGTCATGTACCTGCAAAAGTCAGGGCAAGAC [C,T,G] TGTCCCAGGGCCAAACGCCAGATCAAGCGGGGGGCCGGCCCCGAGGTAAGCGGGGTC CCGGGGCTGGCGGGGTCGAGCGGGCGCACCAAGGGTTGCTCTGTCAGGCCATAGCT TGGCACTGCCGGGGGGCTCTCAGGCTGGCAGGAGAGGGCAGGCCCTCAAGGGGAA AGGGGCTGGACGCCCTGGCCGGTGTGGGCTGGCACGGGGCGGAAGGAAGCGGG ATGCCCGGGGCTTGGGATGGCAGTCCAGGGGGCTCCCGAGAGGGGACGACAG
16266	GAGCCCTGAAACCCCTCTCTGGGCCCTGGGCCCCAGGTGACTGCCAGCCTCAAGGGCTGC

FIGURE 3M

CTCTGTTGAGGAAAGACGACGTACAGGCTACTTCCCGTCCATGTACCTGCAAAAGTC
 GGGCAAGACGTGTCCCAGGCCAACGCCAGATCAAGCGGGGGGCGCCGCCAGGTAA
 GCGGGGGTCCCCGGGCTGGCGGGTCAGCGGGGCGCACCAACGGGTCGCTCTGTCTA
 GGCCATAGCTTGGCAGTGCCGGGCGGGGCTCTAGCCCTGGCAGGAGAGGCAGGACCC
 [C, T]
 ACGGGGGAAAGGGGCTGGACGCGCTGGCCCGGTGTGGGCTGGCACGGGGCGGAAGG
 AAAGCGGGCATGCCGGGGCTTGGGATGGCAGTCCAGGGGGCTCCCCGGAGAGGG
 GGACGACAGACCGAAGGCTGGTGAGGGGCGTGGAAAACGCCAGGCTCTGCTGAGGGC
 AAGGGCTTGTGTCGTACGGGGGAGCCCTTGTCCCAGGGGCTGTGAGACTAC
 CGGCCCCCTACTGCCCTTCACTTCTCGGACCAGGGTGCCTATCTGAGTCCCCTGGGGCAG

 16629 AGCGCGATGCCGGGGCTTGGGATGGCAGTCCAGGGGGCTCCCCGGAGAGGGGG
 ACGACAGACCGAAGGCTGGTGAGGGGCTGGAAAACGCCAGGCTCTGCTGAGGGCA
 GGGCTTGTGTCGTACGGGGGAGCCCTTGTCCCAGGGGCTGTGAGACTAC
 GCCCTACTGCCCTTCACTTCTCGGACCAGGGTGCCTATCTGAGTCCCCTGGGGCAG
 GGGGCCCTGGGCTTGAACGACGCCCCGTCCCGTGGCCAGGTGTCATCCGCAACG
 [C, T]
 GCACAGCATCCACCAAGCGGTGCGGAAGGCCCTAGCCAGGACGCCATGCCGCAACAG
 CGTCGGTTCTGCAACGACGCCAGGCCAGGCGGGGACCGCAGACCCCCGGGAG
 CCCCTCGGTGAGTGCAACGGGAGAGGGCAGGAAGGGCAAGCCCTAGGGGGAGTCAGC
 GGGAGAGGGGGCCAGAGGCAAGGGCAGAGTAGCGGGGGGGGACCAAGAGGGCGGAATCA
 GAGGGAGAGGGGGACTGGAGGCGGGGAGAGGAGGAGCCAGCGTAGGGGGGGAGC

 16642 GGGGGCTTGGGATGGCAGTCCAGGGGGCTCCCGAGAGGGGGACGACAGCCGA
 GGCTGGTGGGGCGTGAAAACGCCAGGCTTGCTGAGGGCAAGGGCTTGTGTC
 GACGGGGCAGCCCTTGTCCCAGGGCTGTGAGACTACCGCCCCCTACTGCC
 CCCCACCTCTCGGACCAGGGTGCCTATCTGAGTCCCCTGGGGCAGGGGCCCTCGG
 CTTGACGACGCCCGTCCCAGGGCTGGCCAGGTGTCATCCGCAACGCGACAGCATCCA
 [C, T]
 CAGCGTCGCGGAAGGCCCTAGCCAGGACGCCATGCCGCAACAGCGTCCGTTCTG
 CAGCAGCGACGCCGCCAGGCGGCCAGGAGGGGAGCCAGAGGCCCCGGAGCCCTCGGTGAG
 TGAGCGGGAGAGGGCAGGAAGGGCAAGCCCTAGGGGGAGTCAGGGAGAGGGGG
 CCAGAGGAGGGCCAGAGTAGCGGGGGGGACCAAGAGGGCGGAATCAGAGGGAGAGGGCG
 GGAAGGGAGGGGGAGAGGAGGAGCCAGCGTAGGGGGGGAGGAGTCCCTAAGAGG

 18537 AAAAAAGTAACCTAGGTGCAAGGGTGTCTCTGTTATTCACTGAGACCGTCCCCGGTTAT
 GAGGTGTACAGAAAGCAAGTATTCACTATGCAACTATTACCGCTCACCTAGCATT
 GAAGCCAGCTGTAGCCTGAAAGCCTTGCCTTGAGGGCAGGTCTTCCCCAAAATGCA
 ACACGAAGGTGCAAAGTGAAGCTGCCAGTCTGCAAAAGATGTAACCTGTCA
 CGAGTGGCAGGGAGAGCTGTCACATTGCGGAAGTGGCTATGTGAGGACGGGGGAGG
 [C, T]
 GGGTCCCTAGAGATAAGAGACAATCATAAAGGGAGATATCAGAGAAAATCGTAAGGGGA
 GCAGATGGTTGCAAGAGAAATAGGCTGACCATGAGGACTGGCAGAAGCTTCAGAAAA
 CCACTGGACGGCTGGCACAGTGGCTTGGCTGTAATCCCAGCACTTGGAGGCTGAC
 GCAGGTGAATCACTTGAGGTCAAGGAGTCCAGACAGCCCTGGCAACATGGTAAACCC
 ATCTCTACAGAAAATATAAAATTAGCCAGGCGTGGTGGACAAGCCTAGAATCCCAGCT

 18589 CCGTTATGAGGTTGTACCAAGAAAGCAAGTATTCACTATGCAACTATTACCGCTCACC
 CTAGCATTGAGGCCAGCTGTAGCCTGAAAGCCTTGCCTTGAGGGCAGGTCTTCCCCCA
 AAATGAGACACGAAGGTGCAAAGTGAAGCTGCCAGTCTGCAAAAGATGTAACCTGTCA
 CGAAGGCCACGAGTGGCAGGGAGAGCTGTCACATTGCGGAAGTGGCTATGTGAGGAC
 GGGGGAGGGGGTCCCTAGAGATAAGAGACAATCATAAAGGGAGATATCAGAGAAAATC
 [G, A]
 TAAGGGGAGCAGATGGTTGCAAGAGAAATAGGCTGACCATGAGGACTGGCAGAAGCTT
 TCAGAAAACACTGGACGGCTGGCACAGTGGCTTGGCTGTAATCCCAGCACTTGG
 AGGCTGACGCGAGGTGAATCACTTGAGGTCAAGGAGTCCAGACAGCCCTGGCAACATGGT
 GAAAACCCATCTACAGAAAATATAAAATTAGCCAGGCGTGGTGGACAAGCCTAGAA
 TCCCAGCTACTGGGAGGCTGAGG

 18720 CGAAGGTGCAAAGTGAAGCTGCCAGTCTGCAAAAGATGTAACCTGTCA
 CGAAGGCCACGAGTGGCAGGGAGAGCTGTCACATTGCGGAAGTGGCTATGTGAGGACGGGGAGGCG
 GTCCCTAGAGATAAGAGACAATCATAAAGGGAGATATCAGAGAAAATGTAAGGGAGC
 AGATGGTTGCAAGAGAAATAGGCTGACCATGAGGACTGGCAGAAGCTTCA
 ACTGGACGGCTGGCACAGTGGCTTGGCTGTAATCCCAGCACTTGGGAGGCTGACGC
 [G, A]

FIGURE 3N

GGTGAATCACTTGAGGTCAAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAACCCCCATC
TCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGCACAAGCCTAGAATCCCAGCTACT
TGGGAGGCTGAGG

18782 TGGCAGGGAGAGCTGTCCCACATTGCGGAAGTGGCTATGTGAGGACGGGGAGGCGGGT
CCCTTAGAGATAAGAGACAATCATAAGGGAGATATCAGAGAAAATCGTAAGGGGAGCAG
ATGGTTGTCAAGAGAAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTCAGAAAACAC
TGGACGGCTGGCACAGTGGCTAGGCCTGTAATCCCAGCACTTTGGGAGGCTGACGCA
GTGAATCACTTGAGGTCAAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAACCCCCATCT
[C, T]
TACAGAAAATATAAAAATTAGCCAGGCGTGGTGCACAAGCCTAGAATCCCAGCTACTTG
GGAGGCTGAGG

18841 TCCCTTAGAGATAAGAGACAATCATAAGGGAGATATCAGAGAAAATCGTAAGGGGAGCA
GATGGTTGTCAAGAGAAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTCAGAAAACCA
CTGGACGGCTGGCACAGTGGCTTAGGCCTGTAATCCCAGCACTTTGGGAGGCTGACGCA
GGTGAATCACTTGAGGTCAAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAACCCCCATC
TCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGCACAAGCCTAGAATCCCAGCTACT
[C, T]
GGGAGGCTGAGG

Chromosome mapping
Chromosome 19

FIGURE 3O